

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:SSSPTA1653HXP

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * Welcome to STN International * * * * *

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America
NEWS 2 "Ask CAS" for self-help around the clock
NEWS 3 JUL 20 Powerful new interactive analysis and visualization software,
STN AnaVist, now available
NEWS 4 AUG 11 Derwent World Patents Index(R) web-based training during
August
NEWS 5 AUG 11 STN AnaVist workshops to be held in North America
NEWS 6 AUG 30 CA/CAPLUS - Increased access to 19th century research documents
NEWS 7 AUG 30 CASREACT - Enhanced with displayable reaction conditions
NEWS 8 SEP 09 ACD predicted properties enhanced in REGISTRY/ZREGISTRY

NEWS EXPRESS JUNE 13 CURRENT WINDOWS VERSION IS V8.0, CURRENT
MACINTOSH VERSION IS V6.0c(ENG) AND V6.0Jc(JP),
AND CURRENT DISCOVER FILE IS DATED 13 JUNE 2005

NEWS HOURS STN Operating Hours Plus Help Desk Availability
NEWS INTER General Internet Information
NEWS LOGIN Welcome Banner and News Items
NEWS PHONE Direct Dial and Telecommunication Network Access to STN
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

All use of STN is subject to the provisions of the STN Customer
agreement. Please note that this agreement limits use to scientific
research. Use for software development or design or implementation
of commercial gateways or other similar uses is prohibited and may
result in loss of user privileges and other penalties.

* * * * * STN Columbus * * * * *

FILE 'HOME' ENTERED AT 13:07:27 ON 20 SEP 2005

=> file medline, uspatful, dgene, embase, wpids, biosis, biotechds
COST IN U.S. DOLLARS SINCE FILE TOTAL
ENTRY SESSION
FULL ESTIMATED COST 0.21 0.21

FILE 'MEDLINE' ENTERED AT 13:08:15 ON 20 SEP 2005

FILE 'USPATFULL' ENTERED AT 13:08:15 ON 20 SEP 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'DGENE' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'EMBASE' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 Elsevier Inc. All rights reserved.

FILE 'WPIDS' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'BIOSIS' ENTERED AT 13:08:15 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'BIOTECHDS' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

=> s (integrating plasmid)
L1 272 (INTEGRATING PLASMID)

=> s candida utilis
L2 5634 CANDIDA UTILIS

=> s l2 and (biotin synthase)
L3 11 L2 AND (BIOTIN SYNTHASE)

=> d l3 ti abs ibib tot

L3 ANSWER 1 OF 11 USPATFULL on STN

TI Methods for monitoring multiple gene expression

AB The present invention relates to methods for monitoring differential expression of a plurality of genes in a first filamentous fungal cell relative to expression of the same genes in one or more second filamentous fungal cells using microarrays containing filamentous fungal expressed sequenced tags. The present invention also relates to filamentous fungal expressed sequenced tags and to computer readable media and substrates containing such expressed sequenced tags for monitoring expression of a plurality of genes in filamentous fungal cells.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2005:140340 USPATFULL

TITLE: Methods for monitoring multiple gene expression

INVENTOR(S): Berka, Randy M., Davis, CA, UNITED STATES

Rey, Michael W., Davis, CA, UNITED STATES

Shuster, Jeffrey R., Davis, CA, UNITED STATES

Kauppinen, Sakari, Smoerum, DENMARK

Clausen, Ib Groth, Hillerod, DENMARK

Olsen, Peter Bjarke, Copenhagen, DENMARK

PATENT ASSIGNEE(S): Novozymes Biotech, Inc., Davis, CA, UNITED STATES (U.S. corporation)

Novozymes A/S, Bagsvaerd, DENMARK (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6902887	B1	20050607
APPLICATION INFO.:	US 2000-533559		20000322 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 1999-273623, filed on 22 Mar 1999, ABANDONED		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	GRANTED		
PRIMARY EXAMINER:	Brusca, John S.		
LEGAL REPRESENTATIVE:	Stames, Robert L.		
NUMBER OF CLAIMS:	8		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	0 Drawing Figure(s); 0 Drawing Page(s)		
LINE COUNT:	18386		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 2 OF 11 USPATFULL on STN

TI Methods for monitoring multiple gene expression

AB The present invention relates to methods for monitoring differential expression of a plurality of genes in a first filamentous fungal cell relative to expression of the same genes in one or more second filamentous fungal cells using microarrays containing filamentous fungal expressed sequenced tags. The present invention also relates to filamentous fungal expressed sequenced tags and to computer readable media and substrates containing such expressed sequenced tags for monitoring expression of a plurality of genes in filamentous fungal cells.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:292283 USPATFULL

TITLE: Methods for monitoring multiple gene expression

INVENTOR(S): Berka, Randy M., Davis, CA, UNITED STATES

Rey, Michael W., Davis, CA, UNITED STATES

Shuster, Jeffrey R., Davis, CA, UNITED STATES

Kauppinen, Sakari, Smoerum, DENMARK

Clausen, Ib Groth, Hillerod, DENMARK

Olsen, Peter Bjarke, Copenhagen, DENMARK

PATENT ASSIGNEE(S): Novozymes Biotech, Inc., Davis, CA (U.S. corporation)
Novozymes A/S, Inc., Bagsvaerd, DENMARK (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004229367	A1	20041118
APPLICATION INFO.:	US 2003-653047	A1	20030829 (10)
RELATED APPLN. INFO.:	Division of Ser. No. US 2000-533559, filed on 22 Mar 2000, PENDING Continuation-in-part of Ser. No. US 1999-273623, filed on 22 Mar 1999, ABANDONED		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	APPLICATION		
LEGAL REPRESENTATIVE:	NOVOZYMES BIOTECH, INC., 1445 DREW AVE, DAVIS, CA, 95616		
NUMBER OF CLAIMS:	8		
EXEMPLARY CLAIM:	1		
LINE COUNT:	19697		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 3 OF 11 USPATFULL on STN

TI Yeast with high biotin-productivity and the preparation method thereof

AB The invention discloses a yeast with high biotin-productivity and the preparation method thereof. The yeast is transformed by an integrated plasmid, which includes a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2003:152918 USPATFULL

TITLE: Yeast with high biotin-productivity and the preparation method thereof

INVENTOR(S): Shiuan, David, Kaohsiung, TAIWAN, PROVINCE OF CHINA

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003104584	A1	20030605
APPLICATION INFO.:	US 2001-752957	A1	20010102 (9)

NUMBER	DATE
--------	------

PRIORITY INFORMATION: TW 2000-89120972 20001007
 DOCUMENT TYPE: Utility
 FILE SEGMENT: APPLICATION
 LEGAL REPRESENTATIVE: BRUCE LONDA, NORRIS, MCLAUGHLIN & MARCUS, P.A., 220
 EAST 42ND STREET, 30TH FLOOR, NEW YORK, NY, 10017
 NUMBER OF CLAIMS: 16
 EXEMPLARY CLAIM: 1
 NUMBER OF DRAWINGS: 13 Drawing Page(s)
 LINE COUNT: 486
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 4 OF 11 USPATFULL on STN

TI Novel Polynucleotides

AB Novel polynucleotides derived from microorganisms belonging to
 coryneform bacteria and fragments thereof, polypeptides encoded by the
 polynucleotides and fragments thereof, polynucleotide arrays comprising
 the polynucleotides and fragments thereof, recording media in which the
 nucleotide sequences of the polynucleotide and fragments thereof have
 been recorded which are readable in a computer, and use of them.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2002:343879 USPATFULL
 TITLE: Novel Polynucleotides
 INVENTOR(S): Nakagawa, Satoshi, Tokyo, JAPAN
 Mizoguchi, Hiroshi, Tokyo, JAPAN
 Ando, Seiko, Tokyo, JAPAN
 Hayashi, Mikiro, Tokyo, JAPAN
 Ochiai, Keiko, Tokyo, JAPAN
 Yokoi, Haruhiko, Tokyo, JAPAN
 Tateishi, Naoko, Tokyo, JAPAN
 Senoh, Akihiro, Tokyo, JAPAN
 Ikeda, Masato, Tokyo, JAPAN
 Ozaki, Akio, Hofu-shi, JAPAN

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002197605	A1	20021226
APPLICATION INFO.:	US 2000-738626	A1	20001218 (9)

	NUMBER	DATE
PRIORITY INFORMATION:	JP 1999-377484	19991216
	JP 2000-159162	20000407
	JP 2000-280988	20000803
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	NIXON & VANDERHYE P.C., 8th Floor, 1100 North Glebe Road, Arlington, VA, 22201	
NUMBER OF CLAIMS:	68	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	4 Drawing Page(s)	
LINE COUNT:	13673	
CAS INDEXING IS AVAILABLE FOR THIS PATENT.		

L3 ANSWER 5 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN

TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.

AN ADA09589 DNA DGENE

AB The present invention relates to a yeast with high biotin-productivity
 and a method for its preparation. The yeast is transformed by an
 integrated plasmid, which includes the edible yeast (*Candida
 utilis*) biotin synthase gene, an assistant

DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09589 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: PCR primer #1 for *S. cerevisiae* BIO2 gene.

L3 ANSWER 6 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
AN ADA09591 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (*Candida utilis*) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a degenerate PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09591 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: Degenerate PCR primer #1 for *C. utilis* BIO2 gene.

L3 ANSWER 7 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
AN ADA09590 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (*Candida utilis*) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09590 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: PCR primer #2 for S. cerevisiae BIO2 gene.

L3 ANSWER 8 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09592 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (**Candida utilis**) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a degenerate PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09592 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: Degenerate PCR primer #2 for C. utilis BIO2 gene.

L3 ANSWER 9 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09588 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (**Candida utilis**) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents C. utilis **biotin synthase** gene.

ACCESSION NUMBER: ADA09588 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: **Candida utilis biotin**

synthase gene.

L3 ANSWER 10 OF 11 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI New integrated plasmid, useful for preparing yeast with high
biotin-productivity used as feed additives, food additives or cosmetics.
AN 2003-567956 [53] WPIDS
AB US2003104584 A UPAB: 20030820

NOVELTY - An integrated plasmid comprising a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

(1) a method for preparing a yeast with high biotin-productivity comprising constructing the integrated plasmid cited above, linearizing the integrated plasmid, transforming the linearized integrated plasmid into a yeast, and recombining the **biotin synthase** gene with the yeast genome; and

(2) a method for producing biotin comprising providing the yeast with high biotin-productivity, culturing the yeast in a nutrient medium, and recovering biotin from the culture broth.

USE - The integrated plasmid is useful for preparing yeast with high biotin-productivity. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics (all claimed).

Dwg.0/13

ACCESSION NUMBER: 2003-567956 [53] WPIDS
DOC. NO. CPI: C2003-153236
TITLE: New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
DERWENT CLASS: C06 D13 D16 D21
INVENTOR(S): SHIUAN, D
PATENT ASSIGNEE(S): (NASC-N) NAT SCI COUNCIL; (SHIU-I) SHIUAN D
COUNTRY COUNT: 2
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2003104584	A1	20030605	(200353)*		20
TW 220675	B1	20040901	(200522)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003104584	A1	US 2001-752957	20010102
TW 220675	B1	TW 2000-120972	20001007

PRIORITY APPLN. INFO: TW 2000-120972 20001007

L3 ANSWER 11 OF 11 BIOTECHDS COPYRIGHT 2005 THE THOMSON CORP. on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics; vector-mediated gene transfer and expression in host cell for strain improvement and feed additive, food additive or cosmetic manufacture
AN 2003-20688 BIOTECHDS
AB DERWENT ABSTRACT:
NOVELTY - An integrated plasmid comprising a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following: (1) a method for preparing a yeast with high biotin-productivity comprising constructing the integrated plasmid cited above, linearizing the integrated plasmid, transforming the linearized integrated plasmid into a yeast, and recombining the **biotin synthase** gene with the yeast genome; and (2) a method for producing biotin comprising providing the yeast with high biotin-productivity, culturing the yeast in a nutrient medium, and recovering biotin from the culture broth.

BIOTECHNOLOGY - Preferred Plasmid: The **biotin synthase** gene of the integrated plasmid is derived from *Saccharomyces cerevisiae* or *Candida utilis*. The **biotin synthase** gene of *C. utilis* comprises a fully defined sequence of 1188 base pairs (bp) given in the specification. The assistant DNA sequence is a *C. utilis* fragment consisting of *NsiI*-*BamHI* 18s rDNA, *URA3* DNA, or *HIS3* DNA. The selection marker is preferably a cycloheximide-resistant gene. The promoter sequence is pL41 promoter of *C. utilis* or *pADH1* promoter of *S. cerevisiae*. The integrated plasmid is: (A) pMCC21 comprising pL41 for the expression of *C. utilis* BIO2 gene encoding for 204 amino acids, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA; (B) pMCC31S or pMCC38S comprising pL41 for the expression of *C. utilis* BIO2 gene encoding for 233 amino acids, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA; (C) pMCC32H or pMCC36H comprising *C. utilis* pL41, *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* *HIS3* gene fragment; or (D) pMCC33U or pMCC35U comprising *C. utilis* pL41, *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* *URA3* gene fragment.

USE - The integrated plasmid is useful for preparing yeast with high biotin-productivity. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics (all claimed).

EXAMPLE - Polymerase chain reaction (PCR) was carried out using the *Candida utilis* BIO2 gene as a template. The product was digested with *NdeI* and *SalI*, and ligated into pMC5 to obtain pMCC11. The *NdeI*-*SalI* fragment of pMCC11 in which the *NdeI* nick was end-repaired into a blunt end by Klenow DNA polymerase (i.e. pL41-BIO2 fragment) was ligated into pMC9 to obtain pMCC15. To fill the lost L41 fragment, the *BamHI*-*EcoRI* fragment of pMC9 (i.e. L41 gene fragment) was end-repaired into a blunt end by Klenow DNA polymerase and then integrated into the *SalI* site in pMCC15 to obtain pMCC21. The plasmid pMCC21 comprises *C. utilis* pL41 for the expression of the *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA. (20 pages)

ACCESSION NUMBER: 2003-20688 BIOTECHDS

TITLE: New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics;
vector-mediated gene transfer and expression in host cell for strain improvement and feed additive, food additive or cosmetic manufacture

AUTHOR: SHIUAN D

PATENT ASSIGNEE: SHIUAN D

PATENT INFO: US 2003104584 5 Jun 2003

APPLICATION INFO: US 2001-752957 2 Jan 2001

PRIORITY INFO: TW 2000-120972 7 Oct 2000; TW 2000-120972 7 Oct 2000

DOCUMENT TYPE: Patent

LANGUAGE: English

OTHER SOURCE: WPI: 2003-567956 [53]

=> s yeast preparation

L4 310 YEAST PREPARATION

=> s 14 and (biotin production)

L5 0 L4 AND (BIOTIN PRODUCTION)

=> s 14 and (biotin)

L6 8 L4 AND (BIOTIN)

=> d 16 ti abs ibib tot

L6 ANSWER 1 OF 8 USPATFULL on STN

TI Whole cell engineering using real-time metabolic flux analysis

AB The invention provides methods for whole cell engineering of new and modified phenotypes by using "on-line" or "real-time" metabolic flux analysis. The invention provides a method for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis by making a modified cell by modifying the genetic composition of a cell and culturing the modified cell to generate a plurality of modified cells and measuring at least one metabolic parameter of the cell by monitoring the cell culture of in real time. The invention also provides articles comprising machine-readable medium including machine-executable instructions and systems, e.g., computer systems, to practice the methods of the invention.

ACCESSION NUMBER: 2005:233489 USPATFULL

TITLE: Whole cell engineering using real-time metabolic flux analysis

INVENTOR(S): Short, Jay M., Rancho Sante Fe, CA, UNITED STATES
Latterich, Martin, Montreal, CANADA
Levin, Michael, San Diego, CA, UNITED STATES
Wei, Jing, San Diego, CA, UNITED STATES
Fu, Pengcheng, Honolulu, HI, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2005202426	A1	20050915
APPLICATION INFO.:	US 2003-491358	A1	20021001 (10)
	WO 2002-US31380		20021001
			20050223 PCT 371 date

	NUMBER	DATE
PRIORITY INFORMATION:	US 2003-326653P	20011001 (60)
	US 2003-326654P	20011001 (60)
	US 2003-326655P	20011001 (60)
	US 2003-337526P	20011109 (60)

DOCUMENT TYPE: Utility

FILE SEGMENT: APPLICATION

LEGAL REPRESENTATIVE: DIVERSA C/O MOFO S.D., 3811 VALLEY CENTER DRIVE, SUITE 500, SAN DIEGO, CA, 92130, US

NUMBER OF CLAIMS: 80

EXEMPLARY CLAIM: 1

NUMBER OF DRAWINGS: 40 Drawing Page(s)

LINE COUNT: 4859

L6 ANSWER 2 OF 8 USPATFULL on STN

TI Screening method for genes of brewing yeast

AB The present invention provides (A) a method for the selection of genes participating in the desired brewing character, which comprises preparing a database compiling the data of the whole genome sequence of industrial yeast, particularly a brewing yeast used for alcoholic beverages; selecting gene participating in a desired brewing character that the brewing yeast specifically possesses; and carrying out

functional analysis of the gene by disruption or overexpression; (B) a DNA array in which oligonucleotide(s) selected based on the data base compiling the data of the whole genome sequences of an industrial yeast, (C) a breeding method for constructing improved cultures achieving the desired brewing character, (D) a method for producing an alcohol or an alcoholic beverage in which productivity and quality are improved using the yeast, (E) a gene which is specific to the improved brewing yeast, and (F) a peptide encoded by the gene.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:334780 USPATFULL
TITLE: Screening method for genes of brewing yeast
INVENTOR(S): Nakao, Yoshihiro, Kyoto-shi, JAPAN
Nakamura, Norihisa, Kyoto-shi, JAPAN
Kodama, Yukiko, Osaka, JAPAN
Fujimura, Tomoko, Osaka, JAPAN
Ashikari, Toshihiko, Osaka, JAPAN
PATENT ASSIGNEE(S): SUNTORY LIMITED, OSAKA, JAPAN (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004265862	A1	20041230
APPLICATION INFO.:	US 2004-791791	A1	20040304 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	JP 2003-57677	20030304
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	BURNS DOANE SWECKER & MATHIS L L P, POST OFFICE BOX 1404, ALEXANDRIA, VA, 22313-1404	
NUMBER OF CLAIMS:	46	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	12 Drawing Page(s)	
LINE COUNT:	3038	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L6 ANSWER 3 OF 8 USPATFULL on STN

TI Whole cell engineering using real-time metabolic flux analysis
AB The invention provides methods for whole cell engineering of new and modified phenotypes by using "on-line" or "real-time" metabolic flux analysis. The invention provides a method for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis by making a modified cell by modifying the genetic composition of a cell and culturing the modified cell to generate a plurality of modified cells and measuring at least one metabolic parameter of the cell by monitoring the cell culture of in real time. The invention also provides articles comprising machine-readable medium including machine-executable instructions and systems, e.g., computer systems, to practice the methods of the invention.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:44982 USPATFULL
TITLE: Whole cell engineering using real-time metabolic flux analysis
INVENTOR(S): Fu, Pengcheng, Honolulu, HI, UNITED STATES
Latterich, Martin, San Diego, CA, UNITED STATES
Levin, Michael, San Diego, CA, UNITED STATES
Wei, Jing, San Diego, CA, UNITED STATES
Short, Jay M., Rancho Santa Fe, CA, UNITED STATES
PATENT ASSIGNEE(S): Diversa Corporation, San Diego, CA (U.S. corporation)

NUMBER	KIND	DATE
--------	------	------

PATENT INFORMATION:	US 2004033975	A1	20040219
APPLICATION INFO.:	US 2002-264989	A1	20021001 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2001-326655P	20011001 (60)
	US 2001-326654P	20011001 (60)
	US 2001-326653P	20011001 (60)
	US 2001-337526P	20011109 (60)

DOCUMENT TYPE: Utility
FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: GREGORY P. EINHORN, Fish & Richardson P.C., Suite 500,
4350 La Jolla Village Drive, San Diego, CA, 92122
NUMBER OF CLAIMS: 228
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 40 Drawing Page(s)
LINE COUNT: 5427
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L6 ANSWER 4 OF 8 USPATFULL on STN
TI Baker's yeast and a method producing it
AB A process is disclosed for the production of baker's yeast which
comprises the fermentation of a baker's yeast strain using a
non-molasses carbon source until a fermentation broth having at least
10%, preferably at least 13%, more preferably at least 16% of dry solids
content is formed, which can be used directly as a cream yeast without
concentration.

ACCESSION NUMBER: 1999:72304 USPATFULL
TITLE: Baker's yeast and a method producing it
INVENTOR(S): Plomp, Pieter Jan Arnoldus Maria, Ex Delft, Netherlands
PATENT ASSIGNEE(S): Gist-brocades, B.V., Netherlands (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 5916609		19990629
APPLICATION INFO.:	US 1997-901312		19970728 (8)

	NUMBER	DATE
PRIORITY INFORMATION:	EP 1996-202119	19960727
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	Granted	
PRIMARY EXAMINER:	Sayala, Chhaya D.	
LEGAL REPRESENTATIVE:	Morrison & Foerster, LLP	
NUMBER OF CLAIMS:	21	
EXEMPLARY CLAIM:	1	
LINE COUNT:	537	

L6 ANSWER 5 OF 8 USPATFULL on STN
TI Production of distillers yeast
AB Distillers yeast is produced by a process which results in more rapid
yeast growth, higher cell concentrations and greatly extended yeast
viability. The process involves propagating yeast on a culture medium of
cooked cereal grain mash which has been subjected to enzymatic action of
glucamylase obtained from fungi of the genus Aspergillus. After
propagation, as a result of substantially longer viability, the yeast
can be stored in the culture medium for up to 3 days or longer with
portions being withdrawn from time to time for inoculating distillery
grain mashes.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 75:10033 USPATFULL
 TITLE: Production of distillers yeast
 INVENTOR(S): Van Lanen, deceased, James M., late of Peoria, IL,
 United States
 Smith, Merritt B., Peoria, IL, United States
 Maisch, Weldon F., Peoria, IL, United States
 PATENT ASSIGNEE(S): Hiram Walker & Sons, Inc., Peoria, IL, United States
 (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 3868307		19750225
APPLICATION INFO.:	US 1973-388359		19730815 (5)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 1971-189678, filed on 15 Oct 1971, now abandoned		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	Granted		
PRIMARY EXAMINER:	Naff, David M.		
LEGAL REPRESENTATIVE:	Kirkland and Ellis		
NUMBER OF CLAIMS:	6		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	2 Drawing Figure(s); 2 Drawing Page(s)		
LINE COUNT:	633		
CAS INDEXING IS AVAILABLE FOR THIS PATENT.			

L6 ANSWER 6 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
 TI Bread-baking yeast preparation - involves adding aryl-thioacetic acid derivs.
 to des thio-**biotin** used in molasses-based nutrient medium for
 yeast cultivation.
 AN 1988-159690 [23] WPIDS
 AB SU 1350172 A UPAB: 19930923
 The bread-baking yeasts are obtd. by multistage cultivation under aeration
 conditions on a nutrient medium containing molasses, mineral salts and
 desthiobiotin. During the stages of production of yeast inoculant and
 commercial yeasts, derivs. of arylthio (sulphonyl) acetic acid are added
 in amount (1-10000)x10 power minus 10% jointly with desthiobiotin.
 The specified derivs. are 2-hydroxyalkyl-ammonium salts of formula
 (I), where (a) R=H, R2=4CH3, n=2 and X=N+H3CH(CH2OH)CH(OH)C6H4NO2-4; (b)
 R1=H, R2=4Cl, n=0 and X=N+H(CH3)2CH2CH2OH; or (c) R1=H, R2=4Cl, n=2, and
 X=N+H(CH2CH2OH)3.
 (I) contain metabolically active sulphur which can be incorporated
 into desthiobiotin and form **biotin**. (I) also act as
 antioxidants by lowering the intensity of peroxide oxidation of lipids of
 biological membranes.
 USE/ADVANTAGE - In the microbiological industry. Increased biomass
 yield is obtd. Bul.41/7.11.87
 0/0

ACCESSION NUMBER: 1988-159690 [23] WPIDS
 DOC. NO. CPI: C1988-071563
 TITLE: Bread-baking yeast preparation - involves adding
 aryl-thioacetic acid derivs. to des thio-**biotin**
 used in molasses-based nutrient medium for yeast
 cultivation.
 DERWENT CLASS: D11 D16 E19
 INVENTOR(S): GUSEVA, S A; LEVKOVSKAY, G G; MIRSKOVA, A N
 PATENT ASSIGNEE(S): (ASII) AS SIBE IRKUT ORG CHEM; (BREA) LENGD BREAD BAKING
 IND
 COUNTRY COUNT: 1
 PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
SU 1350172	A	19871107	(198823)*		3

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
SU 1350172	A	SU 1985-3960649	19851005

PRIORITY APPLN. INFO: SU 1985-3960649 19851005

L6 ANSWER 7 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI Yeast prepared by culturing Hansenula ofunaensis - in medium containing methanol, ethanol and/or glucose as carbon source.
AN 1978-02953A [02] WPIDS
AB JP 52139773 A UPAB: 19930901
The method is characterised by culturing a novel yeast stock Hansenula of unaensis (FERM-P 3494). The yeast stock Hansenula of unaensis can utilise methanol, ethanol and glucose, which are favourable as the starting material for preparing yeast for food use.

The culture is practiced aerobically at 10-41 pref. 30-39 degrees C in the medium containing methanol, ethanol and/or glucose as carbon source at pH 2-9 pref. 3-7 for up to 6 days. As the growth promoter vitamins such as thiamine, **biotin**, etc. and amino acid such as methionine, cysteine, etc. can be applied. After the culture yeast body is separated washed and dried.

ACCESSION NUMBER: 1978-02953A [02] WPIDS
TITLE: Yeast prepared by culturing Hansenula ofunaensis - in medium containing methanol, ethanol and/or glucose as carbon source.
DERWENT CLASS: C03 D13 D16
PATENT ASSIGNEE(S): (MITK) MITSUI TOATSU CHEM INC
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
JP 52139773	A	19771121	(197802)*		
JP 56012109	B	19810318	(198115)		

PRIORITY APPLN. INFO: JP 1976-44940 19760422

L6 ANSWER 8 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI Yeast preparation using Torulopsis methanothermophilus - in culture medium containing methanol as carbon source and **biotin**.
AN 1977-19283Y [11] WPIDS
AB JP 52015882 A UPAB: 19930901

Method is effected by inoculating a stock belonging to Torulopsis methanothermophilus which can utilize methanol, into a culture medium containing methanol as carbon source and **biotin**, culturing it for multiplication and separating the yeast body.

The stock can multiply at high temperature and low pH and is considered a novel stock. The stock is trusted to Institute of Microbial Industry with the Trust Number 3158. As teh carbon source crude methanol which contains small amount of lower aldehydes, lower alcohols, etc. can be used. With methanol concentrate >4% the growth of the yest is inhibited, so methanol is added to the culture medium little by little.

ACCESSION NUMBER: 1977-19283Y [11] WPIDS
TITLE: Yeast preparation using Torulopsis methanothermophilus - in culture medium containing methanol as carbon source and **biotin**.
DERWENT CLASS: D16

PATENT ASSIGNEE(S): (MITP) MITSUBISHI PETROCHEMICAL CO LTD
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
JP 52015882	A	19770205	(197711)*		

PRIORITY APPLN. INFO: JP 1975-91807 19750728

=> d his

(FILE 'HOME' ENTERED AT 13:07:27 ON 20 SEP 2005)

FILE 'MEDLINE, USPATFULL, DGENE, EMBASE, WPIDS, BIOSIS, BIOTECHDS'
ENTERED AT 13:08:15 ON 20 SEP 2005

L1 272 S (INTEGRATING PLASMID)
L2 5634 S CANDIDA UTILIS
L3 11 S L2 AND (BIOTIN SYNTHASE)
L4 310 S YEAST PREPARATION
L5 0 S L4 AND (BIOTIN PRODUCTION)
L6 8 S L4 AND (BIOTIN)

=> s l1 and l3
L7 0 L1 AND L3

=> s l2 and l1
L8 2 L2 AND L1

=> d l8 ti abs ibib tot

L8 ANSWER 1 OF 2 USPATFULL on STN
TI Production of farnesol and geranylgeraniol
AB The invention provides a biological method of producing farnesol or geranylgeraniol.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:144592 USPATFULL
TITLE: Production of farnesol and geranylgeraniol
INVENTOR(S): Millis, James R., Kohler, WI, UNITED STATES
Maurina-Brunker, Julie, Appleton, WI, UNITED STATES
McMullin, Thomas W., Manitowoc, WI, UNITED STATES
PATENT ASSIGNEE(S): Arkion Life Sciences LLC. (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004110257	A1	20040610
APPLICATION INFO.:	US 2003-719196	A1	20031120 (10)
RELATED APPLN. INFO.:	Continuation of Ser. No. US 2001-909558, filed on 20 Jul 2001, GRANTED, Pat. No. US 6689593 Continuation of Ser. No. US 1999-350275, filed on 6 Jul 1999, GRANTED, Pat. No. US 6531303		

	NUMBER	DATE
PRIORITY INFORMATION:	US 1998-91964P	19980706 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	SHERIDAN ROSS PC, 1560 BROADWAY, SUITE 1200, DENVER, CO, 80202	

NUMBER OF CLAIMS: 17
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 11 Drawing Page(s)
LINE COUNT: 3450
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L8 ANSWER 2 OF 2 USPATFULL on STN
TI Manufacture of five-carbon sugars and sugar alcohols
AB The invention relates to the methods of manufacturing five-carbon sugars and sugar alcohols as well as other compounds derived from pentose-phosphate pathway from readily available substrates such a hexoses using metabolically engineered microbial hosts.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2003:99684 USPATFULL
TITLE: Manufacture of five-carbon sugars and sugar alcohols
INVENTOR(S): Miasnikov, Andrei, Kantvik, FINLAND
Ojamo, Heikki, Kirkkonummi, FINLAND
Povelainen, Mira, Espoo, FINLAND
Gros, Hakan, Kantvik, FINLAND
Toivari, Mervi, Espoo, FINLAND
Richard, Peter, Helsinki, FINLAND
Ruuhonen, Laura, Helsinki, FINLAND
Koivuranta, Kari, Helsinki, FINLAND
Londesborough, John, Helsinki, FINLAND
Aristidou, Aristos, Espoo, FINLAND
Penttila, Merja, Helsinki, FINLAND
Plazanet-Menut, Claire, Paris, FRANCE
Deutscher, Josef, Fontenay Le Fleury, FRANCE
PATENT ASSIGNEE(S): Xyrofin Oy (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003068791	A1	20030410
APPLICATION INFO.:	US 2001-908744	A1	20010720 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 2000-488581, filed on 21 Jan 2000, ABANDONED Continuation-in-part of Ser. No. US 1997-790585, filed on 29 Jan 1997, PENDING Continuation of Ser. No. US 1995-368395, filed on 3 Jan 1995, GRANTED, Pat. No. US 5631150 Continuation of Ser. No. US 1993-110672, filed on 24 Aug 1993, ABANDONED Continuation-in-part of Ser. No. US 1992-973325, filed on 5 Nov 1992, ABANDONED Continuation-in-part of Ser. No. WO 2001-FI51, filed on 22 Jan 2001, UNKNOWN		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	APPLICATION		
LEGAL REPRESENTATIVE:	STERNE, KESSLER, GOLDSTEIN & FOX PLLC, 1100 NEW YORK AVENUE, N.W., SUITE 600, WASHINGTON, DC, 20005-3934		
NUMBER OF CLAIMS:	166		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	25 Drawing Page(s)		
LINE COUNT:	5422		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

=> file scisearch, biosis, medline, uspatful, wpids, biotechds, embase, dgene
COST IN U.S. DOLLARS

	SINCE FILE	TOTAL
	ENTRY	SESSION
FULL ESTIMATED COST	86.18	86.39

FILE 'SCISEARCH' ENTERED AT 13:16:26 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'BIOSIS' ENTERED AT 13:16:26 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'MEDLINE' ENTERED AT 13:16:26 ON 20 SEP 2005

FILE 'USPATFULL' ENTERED AT 13:16:26 ON 20 SEP 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'WPIDS' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'BIOTECHDS' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'EMBASE' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 Elsevier Inc. All rights reserved.

FILE 'DGENE' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

=> e clemen, r/au

E1	4	CLEMEN U/AU
E2	1	CLEMEN Y/AU
E3	0 -->	CLEMEN, R/AU
E4	1	CLEMENA G/AU
E5	8	CLEMENA G G/AU
E6	8	CLEMENA W/AU
E7	1	CLEMENA WENDY/AU
E8	1	CLEMENACON G/AU
E9	20	CLEMENCE A/AU
E10	5	CLEMENCE A E/AU
E11	13	CLEMENCE A J/AU
E12	4	CLEMENCE B/AU

=> e gerwig, g/au

E1	2	GERWIG W H/AU
E2	55	GERWIG W H JR/AU
E3	0 -->	GERWIG, G/AU
E4	34	GERWIN B/AU
E5	5	GERWIN B D/AU
E6	310	GERWIN B I/AU
E7	1	GERWIN B L/AU
E8	1	GERWIN BIRGIT/AU
E9	1	GERWIN BRENDA/AU
E10	18	GERWIN BRENDA I/AU
E11	19	GERWIN C/AU
E12	19	GERWIN C M/AU

=> e haess, j/au

E1	1	HAESS K/AU
E2	3	HAESS M/AU
E3	0 -->	HAESS, J/AU
E4	3	HAESSE W/AU
E5	1	HAESSEL DE MENENDEZ G/AU
E6	1	HAESSEL DE MENENDEZ G G/AU
E7	1	HAESSEL DE MENENDEZ GABRIELA G/AU
E8	1	HAESSEL U/AU
E9	7	HAESSEL W/AU
E10	1	HAESSELBA U/AU
E11	1	HAESSELBARTH/AU
E12	1	HAESSELBARTH A/AU

=> e fleischer, b/au

E1	15	FLEISCHER WOLFGANG/AU
E2	2	FLEISCHER Z/AU
E3	0 -->	FLEISCHER, B/AU
E4	1	FLEISCHERDJOLETO O/AU
E5	1	FLEISCHERE ARTHUR C/AU
E6	1	FLEISCHERF/AU
E7	1	FLEISCHERH R/AU
E8	2	FLEISCHERKURTZ G/AU
E9	2	FLEISCHERLAMBROPOULOS E/AU
E10	6	FLEISCHERLAMBROPOULOS H/AU
E11	1	FLEISCHERN R V/AU
E12	1	FLEISCHEROVA J/AU

=> e schwarz, e/au

E1	4	SCHWARZ ZS/AU
E2	1	SCHWARZ ZVI/AU
E3	0 -->	SCHWARZ, E/AU
E4	2	SCHWARZA HG/AU
E5	1	SCHWARZA M J/AU
E6	1	SCHWARZABAUER J E/AU
E7	2	SCHWARZACH R/AU
E8	6	SCHWARZACH T/AU
E9	1	SCHWARZACH T T/AU
E10	2	SCHWARZACHER G/AU
E11	1	SCHWARZACHER G M/AU
E12	1	SCHWARZACHER GERTRUDE MARIA/AU

=> e sigal, l/au

E1	97	SIGAL ZAFRANI B/AU
E2	29	SIGAL ZAFRANI BRIGITTE/AU
E3	0 -->	SIGAL, L/AU
E4	11	SIGALA B/AU
E5	8	SIGALA BARBARA/AU
E6	18	SIGALA C/AU
E7	2	SIGALA DE C C/AU
E8	1	SIGALA DE CAMPOS C/AU
E9	79	SIGALA F/AU
E10	2	SIGALA FILIS/AU
E11	13	SIGALA FRAGISKA/AU
E12	1	SIGALA FRANCHESKA/AU

Refine Search

Search Results -

Terms	Documents
L7 and L1	1

Database:

US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search:

L8

Refine Search

Recall Text

Clear

Interrupt

Search History

DATE: Tuesday, September 20, 2005 [Printable Copy](#) [Create Case](#)

Set Name Query

side by side

Hit Count Set Name

result set

DB=PGPB,USPT; PLUR=YES; OP=OR

<u>L8</u>	L7 and l1	1	<u>L8</u>
<u>L7</u>	L6 and (saccharomyces cerevisae)	23150	<u>L7</u>
<u>L6</u>	L5 and (cycloheximide-resistance gene)	73128	<u>L6</u>
<u>L5</u>	L4 and (yeast preparation)	76057	<u>L5</u>
<u>L4</u>	L2 and integrating plasmid	82646	<u>L4</u>
<u>L3</u>	L2 and l1	3	<u>L3</u>
<u>L2</u>	candida utilis biotin synthase	79111	<u>L2</u>
<u>L1</u>	shiuan.in.	76	<u>L1</u>

END OF SEARCH HISTORY

Hit List

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Search Results - Record(s) 1 through 1 of 1 returned.

☐ 1. Document ID: US 20030104584 A1

Using default format because multiple data bases are involved.

L8: Entry 1 of 1

File: PGPB

Jun 5, 2003

PGPUB-DOCUMENT-NUMBER: 20030104584

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030104584 A1

TITLE: Yeast with high biotin-productivity and the preparation method thereof

PUBLICATION-DATE: June 5, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
<u>Shiuan</u> , David	Kaohsiung		TW	

US-CL-CURRENT: 435/119; 435/254.2, 435/320.1, 435/483

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	KWIC	Draw Desc	Imgs
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	------

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Terms	Documents
L7 and L1	1

Display Format:

[Change Format](#)[Previous Page](#)[Next Page](#)[Go to Doc#](#)

Hit List

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Search Results - Record(s) 1 through 3 of 3 returned.

☐ 1. Document ID: US 20030104584 A1

Using default format because multiple data bases are involved.

L3: Entry 1 of 3

File: PGPB

Jun 5, 2003

PGPUB-DOCUMENT-NUMBER: 20030104584

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030104584 A1

TITLE: Yeast with high biotin-productivity and the preparation method thereof

PUBLICATION-DATE: June 5, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
<u>Shiuan</u> , David	Kaohsiung		TW	

US-CL-CURRENT: 435/119; 435/254.2, 435/320.1, 435/483

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	KWIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	-----

☐ 2. Document ID: US 20030003549 A1

L3: Entry 2 of 3

File: PGPB

Jan 2, 2003

PGPUB-DOCUMENT-NUMBER: 20030003549

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030003549 A1

TITLE: DETECTION OF HUMAN HEPATITIS B VIRUS SURFACE ANTIGEN MUTANTS BY SPECIFIC AMPLIFICATION AND ITS APPLICATION ON GENE CHIP

PUBLICATION-DATE: January 2, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
OON, CHONG-JIN	SINGAPORE		SG	
CHEN, WEI-NING	SINGAPORE		SG	
LEONG, AI-LIN	SINGAPORE		SG	
KOH, <u>SHIUAN</u>	SINGAPORE		SG	

US-CL-CURRENT: 435/110

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	KWIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	-----

☐ 3. Document ID: US 6759193 B2

L3: Entry 3 of 3

File: USPT

Jul 6, 2004

TITLE: Detection of human hepatitis B virus surface antigen mutants by specific amplification and its application on gene chip

DATE-ISSUED: July 6, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Oon; Chong-Jin	Singapore			SG
Chen; Wei-Ning	Singapore			SG
Leong; Ai-Lin	Singapore			SG
Koh; Shiuan	Singapore			SG

US-CL-CURRENT: 435/6; 435/5, 435/91.2, 536/23.72, 536/24.3, 536/24.33

Full	Title	Citation	Front	Review	Classification	Date	Reference			Claims	KMIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	--	--	--------	------	-----------	-----

Clear	Generate Collection	Print	Fwd Refs	Bkwd Refs	Generate OACS
-------	---------------------	-------	----------	-----------	---------------

Terms	Documents
L2 and L1	3

Display Format:

[Previous Page](#) [Next Page](#) [Go to Doc#](#)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 13:24:44 ; Search time 5408 Seconds
(without alignments)
10644.378 Million cell updates/sec

Title: US-09-752-957F-1
Perfect score: 1188
Sequence: 1 atgtcgttatattgactgc.....aggttgcatcggtgttga ll88

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	5278	8 AF212161	AF212161 Candida u
C	2	519.4	43.7	4068	8 SCYGR286C
	3	519.4	43.7	35400	8 SCCHIIRA
	4	475.8	40.1	110000	8 CR382121_05
	5	473.6	39.9	1847	8 SCBIO2GNA
	6	458.8	38.6	110000	8 CR382130_18
C	7	451.6	38.0	110000	8 CR382134_04
	8	414.4	34.9	1263	6 AR549314
	9	343.2	28.9	40005	8 SPC320
C	10	343.2	28.9	40880	8 SPCCL235
	11	332.2	28.0	847	11 CMS06K9X
	12	328.2	27.6	3083	6 AR075303
	13	328.2	27.6	3083	6 I66541
	14	328.2	27.6	3083	6 AR306523
C	15	328.2	27.6	3083	8 SCZU01
	16	327.2	27.5	1091	8 SPA224930
	17	318.2	26.8	1447	8 AK101236
	18	307.2	25.9	1440	8 AK059296
	19	304	25.6	1137	6 CQ805820

20	304	25.6	1137	6	AX505884	Sequence
21	304	25.6	1296	8	ATU31806	L34413 Arabidopsis
22	304	25.6	1332	8	ATU31806	ATU31806 Arabidopsis
23	304	25.6	1351	6	AR029503	Sequence
24	304	25.6	1351	6	AR034920	Sequence
25	302.4	25.5	1307	8	BT003848	AT003848 Arabidops
26	300.8	25.3	1342	8	AY088152	Arabidops
27	292.6	24.6	1041	6	AI1523	E.coli B10
28	255.4	21.5	343550	1	AP003587	Nostoc sp
29	243.4	20.5	1032	8	BT009521	Triticum
30	236	19.9	1017	6	CR320148	Sequence
31	231.2	19.5	348672	1	CR378670	Photobact
32	229.4	19.3	110000	1	AE017354_16	Continuation (17 o
33	229.4	19.3	110000	2	AP006494_5	Continuation (6 of
34	228.8	19.3	1421	1	AF239257	Acinetoba
35	228.8	19.3	349174	1	AB063522	AF239257 Wigglewo
36	226.4	19.1	210710	6	AX770902	Sequence
37	226.4	19.1	349071	1	EX571864	AX770902 Sequence
38	224.6	18.9	110000	1	CR628337_17	BX571864 Photobab
39	223.6	18.8	5729	6	CQ872877	Continuation (18 o
40	223.6	18.8	11107	1	U32783	Continuation (18 o
41	223.6	18.8	349980	6	CQ873072	CQ872877 Sequence
42	223	18.8	110000	1	CR628336_15	CQ873072 Sequence
43	215.2	18.1	326050	1	EX321864	Continuation (16 o
44	215	18.1	348178	1	EX572092	BX321864 Nitroso
45	213.8	18.0	4000	1	AF250768	EX572092 Prochlo
46	213.8	18.0	4000	1	AF250768	AF250768 Uncultu

ALIGNMENTS

RESULT 1
AF212161
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AF212161
Candida utilis biotin synthase (Bio2) gene, complete cds.
AF212161.1 GI:11596106
Pichia jadinii
Pichia jadinii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
1 (bases 1 to 5278)
Hong, Y.R. and Shuan, D.
Cloning and sequencing of Candida utilis biotin synthase
Unpublished
2 (bases 1 to 5278)
Hong, Y.R. and Shuan, D.
Direct Submission
Submitted (05-DEC-1999) Graduate Institute of Biochemistry, No.
100, Shih-Chuan 1st Road, Kaohsiung, Taiwan

1..5278
/organism="Pichia jadinii"
/mol_type="genomic DNA"
/db_xref="taxon:4903"
<3915..>5102
/gene="Bio2"
<3915..>5102
/gene="Bio2"
/product="biotin synthase"
3915..5102
/gene="Bio2"
/codon_start=1
/product="biotin synthase long form"
/protein_id="AAG38483.1"
/db_xref="GI:11596107"
/translation="MSFILTISRPAI...STSRVASKATLATGATAAEITLDFVTEQM
EVASQEKPNPLEYALSVKTPVNTWKEIKAIYDTPLMDLHYAQVHREFQOPSE
VGLCTLNKTKGCTEDCKYCAQSORVNTGVKAERLIQVDEVIAEAKAKANGSTRFC
MGAWREKGRKSNLKKIEMITAVHDIGMESCVTLGVMDKQDQATLKSAGUTANHN
IDTYKHPYKIVSTRSFDRLKTFKNQVSGLKGCTGILGEGFGEVRSVPLTYLAT
MDQHPESLPINRLVPIKGTPMYEVKNQVEVDIVRTIATRLVMTPTSIIRLAAGRY

TMKBAEQVMCFMAGCNAIFTGKMLTTMCNGWDEDKAMLAkWGLKPMBSFKYKPREVA
FGA"
4044..5102
/gene="Bio2"
/codon_start=1
/product="biotin synthase short form"
/protein_id="AAG38484.1"
/db_xref="GI:11596108"
/translation="MEEVASQEKKNPLEYALSVKTPVNTWKBEIKAIYDPLMDLM
HYAQVHRRFOOPSEVQLTLMNITGGCTDECKYCAOSQRYNTGVKABRIIOQDEVI
EAAKEAKANGSTRFCMGAAWEMKRNKLNKKIEMITAVHDLGMSCVTLIGWMDKDQ
ATELKSAGLTYLNNIDITYKEHPKVI STRSPFDRLKTFKNVQSGSLCAKGLGLG
ETQEDRVFLVTLTMDQHPESLPINRLVPIKGTPMYEEVKNKQVEVDIEIVRTIATAR
LVMPISILRLAAGRVTMEAEQVMCFMAGCNAI FTGKMLTTMCNGWDEDKAMLAkWGL
LKPMBSFKYKPREVAFGA"

ORIGIN

Query Match 100.0%; Score 1188; DB 8; Length 5278;
Best Local Similarity 100.0%; Pred. No. 1.5e-308;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCGTTTATATGACTGCTATTAGTCGTCGATGCTCTTCCACTTCTAGAGTAGCT 3974
Db 3915 ATGTCGTTTATATGACTGCTATTAGTCGTCGATGCTCTTCCACTTCTAGAGTAGCT 3974

Qy 61 TCTAGGCTACTCTGCAACAGCTGCTACTGCTGCTGCGAGATCTTGGAGATGCTTTC 120
Db 3975 TCTAGGCTACTCTGCAACAGCTGCTACTGCTGCTGCGAGATCTTGGAGATGCTTTC 4034

Qy 121 ACGGAACAAATGGAAGAGTGGCTTTCACAGGAGAAAGACCAACCCATTGGAATATGCA 180
Db 4035 ACGGAACAAATGGAAGAGTGGCTTTCACAGGAGAAAGACCAACCCATTGGAATATGCA 4094

Qy 181 TTGTCAGTGAAGACACCACTGCAACAGCTGCAACAAAGAAAGAAATTAAGCTTATATGAC 240
Db 4095 TTGTCAGTGAAGACACCACTGCAACAGCTGCAACAAAGAAAGAAATTAAGCTTATATGAC 4154

Qy 241 ACACCACTCATGAGCTTGTATGATGCTGCTGAGTGCACACAGAGAGGTTCCAAACACT 300
Db 4155 ACACCACTCATGAGCTTGTATGATGCTGCTGAGTGCACACAGAGAGGTTCCAAACACT 4214

Qy 301 TCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGTACCGAGGACTGT 360
Db 4215 TCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGTACCGAGGACTGT 4274

Qy 361 AAGTACTGTGCCAATCAGCGTTTACACACTGTGTCAAGCTGCAAGGAATCATCCAA 420
Db 4275 AAGTACTGTGCCAATCAGCGTTTACACACTGTGTCAAGCTGCAAGGAATCATCCAA 4334

Qy 421 GTTGATGAGGTGATTTGAAGCTGCAAGGAGGCAAGGCAATGGATCTACAAGGTTCTGT 480
Db 4335 GTTGATGAGGTGATTTGAAGCTGCAAGGAGGCAAGGCAATGGATCTACAAGGTTCTGT 4394

Qy 481 ATGGGTGCTGCTGGAGAGAGATGAAGGTAGAAAGTCAAACTTGAAGAAAATCAAGAG 540
Db 4395 ATGGGTGCTGCTGGAGAGAGATGAAGGTAGAAAGTCAAACTTGAAGAAAATCAAGAG 4454

Qy 541 ATGATCACTGCTGTCCATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAATGGTTGAT 600
Db 4455 ATGATCACTGCTGTCCATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAATGGTTGAT 4514

Qy 601 AAAGACCAAGCCACTGAATTTGAAGTGTGCTGGGTTGACGGCTGCAACCAATCAATGAT 660
Db 4515 AAAGACCAAGCCACTGAATTTGAAGTGTGCTGGGTTGACGGCTGCAACCAATCAATGAT 4574

Qy 661 ACTTACAAGGAACACTATATCAAGGTGATCTCCCAAGAGAGCTTTTGATGATAGATTGAAA 720
Db 4575 ACTTACAAGGAACACTATATCAAGGTGATCTCCCAAGAGAGCTTTTGATGATAGATTGAAA 4634

Qy 721 ACATTCAAAAAGCTTCAAGGATCTGGAATTAAGGCAATGCAAGGTGGTATTTCTGCTTT 780
Db 4635 ACATTCAAAAAGCTTCAAGGATCTGGAATTAAGGCAATGCAAGGTGGTATTTCTGCTTT 4694

781 GGTGAGACCCAGAGGACCGTGATCTTTCTCTACACCTTGGCCACAAATGGATCAGCAT 840
Db 4695 GGTGAGACCCAGAGGACCGTGATCTTTCTCTACACCTTGGCCACAAATGGATCAGCAT 4754

Qy 841 CCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAGGGCAGCCCAATGTATGAAGAA 900
Db 4755 CCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAGGGCAGCCCAATGTATGAAGAA 4814

Qy 901 GTTAAAGAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTG 960
Db 4815 GTTAAAGAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTG 4874

Qy 961 GTCATGCCAACGTCCTATTATCAGATTGGCTCAGGAAGATATACAATGAAGAGGAGCAGAA 1020
Db 4875 GTCATGCCAACGTCCTATTATCAGATTGGCTCAGGAAGATATACAATGAAGAGGAGCAGAA 4934

Qy 1021 CAGTGATGCTCTTCCATGGCTGGTTGTAATGCCATCTTCACAGGTAAAGAAATGCTCACA 1080
Db 4935 CAGTGATGCTCTTCCATGGCTGGTTGTAATGCCATCTTCACAGGTAAAGAAATGCTCACA 4994

Qy 1081 ACAATGTGTAACGGCTGGGATGAGGATAAAGCCCATTTGGCTAAATGGGGTCTGAAACCA 1140
Db 4995 ACAATGTGTAACGGCTGGGATGAGGATAAAGCCCATTTGGCTAAATGGGGTCTGAAACCA 5054

Qy 1141 ATGGAGAGTTTCAAATACAAACCAAGGAGGTTGCAATTCGGTGGCTTGA 1188
Db 5055 ATGGAGAGTTTCAAATACAAACCAAGGAGGTTGCAATTCGGTGGCTTGA 5102

RESULT 2
SCYGR286C/c 4068 bp DNA linear PLN 11-AUG-1997
LOCUS S.cerevisiae chromosome VII reading frame ORF YGR286C.
DEFINITION
ACCESSION Z73071 Y13135
VERSION Z73071.1 GI:1323522
KEYWORDS
SOURCE
ORGANISM
Saccharomyces cerevisiae (baker's yeast)
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
1 (bases 1 to 4068)
Voet, M. and Volckaert, G.
JOURNAL
REFERENCE 2 (bases 1 to 4068)
MIPS.
AUTHORS
Direct Submission
TITLE
Submitted (14-MAY-1996) Data collected by MIPS on behalf of the
European yeast chromosome VII sequencing project. MIPS at the
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
FEATURES
Location/Qualifiers
1..4068
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/chromosome="VII"
complement (660..1787)
/gene="BIO2"
complement (660..1787)
/gene="BIO2"
/note="ORF YGR286C"
/codon_start=1
/protein_id="CAA97318.1"
/db_xref="GI:1323523"
/db_xref="GOA:P32451"
/db_xref="SGD:S0003518"
/translation="MMSTIYRHLSTARPALTKYATNAAVKSTTASSEASTLGALOYAL
SLDSPSHSKYQLKIEYHTPLLELTHAAQLQHRKWHDPFKVOICLTLMNISKGSSED
CKYCAQSSRNDTGLKAEKMWKVDIEKEAEKRNKSTFRFCGLAARWDMKRGKSAKMR
IQEMVTKVNDMGLTCTVLTGMDVDOQAQKADAGLTAVNNHIDTSRREHYSKVITRTY
DDRLOTINKVOESGIKACTGGLGISEDDHIGFIYTLNSMSPHPESLPINRLVAIK
GTPMAELADPKSKKLOPDEILRTIATARIIVMPKAILNLAAGRYTMKTEQFVCPMAG

CNSIFTGKKMLTTMCNGWDEPKAMLAKWGLOPMEAFKYDRS"

```

RESULT 3
LOCUS SCCHVIIRA 35400 bp DNA linear PLN 14-APR-1997
DEFINITION S.cerevisiae DNA for fragment from chromosome VII.
ACCESSION X94332
VERSION X94332.1 GI:1262143
KEYWORDS
SOURCE
ORGANISM
Saccharomyces cerevisiae (baker's yeast)
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
REFERENCE
1 Volckaert,G., Voet,M. and Robben,J.
Sequence analysis of a near-subtelomeric 35.4 kb DNA segment on the
right arm of chromosome VII from Saccharomyces cerevisiae carrying
the MAL1 locus reveals 15 complete open reading frames, including
ZOO1, BGL2 and BIO2 genes and an ABC transporter gene
Yeast 13 (3), 251-259 (1997)
JOURNAL MEDLINE 97245295
PUBMED 9090054
REFERENCE
2 (bases 1 to 35400)
Volckaert,G.
Direct Submission
Submitted (14-DEC-1995) G. Volckaert, Laboratory of Gene
Technology, Catholic University of Leuven, Willem de Croylaan 42,
B-3001 Leuven, BELGIUM
FEATURES
Location/Qualifiers
1..35400
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="FY1679"
/db_xref="taxon:4932"
/chromosome="VII"
/map="right arm"
/clone="cosmid pEGH452"
/clone_lib="H.tetelin (Louvain-la-Neuve, Belgium)"
ORIGIN
Query Match 43.7%; Score 519.4; DB 8; Length 35400;
Best Local Similarity 70.9%; Pred. No. 1.4e-128;
Matches 707; Conservative 0; Mismatches 281; Indels 9; Gaps 1;
Qy 173 AATATGCATTGTTCAGTGAAGACACCACTGACCACTCGGACCAAGAGAATAATTAAGCTA 232
Db 15763 AATATGCATTGTCTTTAGATGAACCAAGTCATTCTGTGGACAAAATCGCAATTAAGAAGAA 15822
Qy 233 TATATGACACACACTCATGGACTTGATGCATCTGCTCAGGTGCACACAGCAAGGTTCC 292
Db 15823 TTTATCATACCCCACTGCTCGAACTTACTCATGCGACGACAAATTGCGACGACAGAAAAGTGC 15882
Qy 293 AACAACTTCAGAGGTTCAATTGTGCACCTCTTATGAATATCAAACTTGTTGTTGTACCG 352
Db 15883 ACGATCCACCAAGTGCAATTGTGCACATGATGAACATCAATCTGTTGTTGTTCTG 15942
Qy 353 AGAAGCTGAAGTACTGTGCCCAATCAACAGCGTTACAACACTGGTGTCAAGGCTGAAGAA 412
Db 15943 AGGACTGTAAGTATTGTGGCGAGTCTTCGAGAAAGCATACCCGCTCTAAGGCTGAGAAA 16002
Qy 413 TCATCCAAGTTGATGAGGTGATTGAAGCTGCAAGAGGCGCAAGGCCAATGGATCTCAAA 472
Db 16003 TGGTTAAAGTGGATGAAGTGATTAAAGGCGAAGAGGCCAAAAGAACCGGATCTCTA 16062
Qy 473 GGTCTTGATGGGTGCTGCTTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTTGAAGAAA 532
Db 16063 GATTCGCTAGTGCTGCATGGAGAGACATGAAAGGTCGTAATACGCCATGAAAGAA 16122
Qy 533 TCAAGAGATGATCACTGCTGTCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGAA 592
Db 16123 TTCAAGAAATGGTGACCAAGTGAATGATATGGGGCTAGAAACGTTGTTACTTTAGTGA 16182

```

ORIGIN	Query Match	43.7%;	Score 519.4;	DB 8;	Length 4068;
	Best Local Similarity	70.9%;	Pred. No. 1.5e-128;		
	Matches 707;	Conservative 0;	Mismatches 281;	Indels 9;	Gaps 1;
Qy	173	AATATGCATTGTCACTGAAGACACACAGTCACACCTGGACCAAGAAGAAATTTAAAGCTA	232		
Db	1666	AATATGCATTGTCTTTAGTAGAACCAAGTCATTCTGTGACAAAATTCGAATTTAAAGAAA	1607		
Qy	233	TATATGACACACCACTCATGTGACTTGTAGTCATATGTTCAGGTGCAACACACAGAAGTTTCC	292		
Db	1606	TTTATCATACCCACTGCTCGAATTACTCATGACGACAAATTCGACGCAAGAAGTGGC	1547		
Qy	293	AACAACTTCAGAGGTTCAAATGTGCACTCTTATGAATATCAAAAATCGTGTGTTGATCCG	352		
Db	1546	ACGATCCAAACCAAGTGCAAATGTGCACATTTGATGAACATCAAAATCTGGTGTGTTCTG	1487		
Qy	353	AGGACTGTAACTACTGTGCCCAATCACAGCGTTACACACTGTGTTCAGGCTCAAGAA	4112		
Db	1486	AGGACTGTAACTACTGTGCCGAGTCTTCGAGAAAACGATACCGGTCTTAAAGGCTCAGAAA	1427		
Qy	413	TCATCCAAATGTGAGGTGATTGAAGCTGCAAAAGGAGGCAAAAGCCAAATGCAATCTCAAA	472		
Db	1426	TGGTTAAAGTGGATGAAGTGAATTAAGAGGCGAAGAGGCCCAAGAAGAACGGATCTACTA	1367		
Qy	473	GGTCTGTATGGGTGCTGCTTTGGAGAGAGATGAAGGTGAGAAAGTCAAACTTTGAAGAAA	532		
Db	1366	GAITCTGCTAGGTGCTGCATCGGAGAGACATGAAGAGTTCGTAATCAGCCATGAAGAA	1307		
Qy	533	TCAAGAGATGATCACTGCTCTCCATGACCTTGGAAATGGAGAGTTGTTCACCTCGGAA	592		
Db	1306	TTCAGGAATGGTGACCAAAAGTGAATGATATGGGGCTAGAAAACGTGTGTACTTTTAGGTA	1247		
Qy	593	TGGTTGATAAAGACCAAGCCACTGAAATTTGAAAGTGTGTGGGTTTCACGCGCGTCAACCATTA	652		
Db	1246	TGGTTGATCAAGATCAAGCAAGCAATTTGAAAGATGCAAGTTTGACTGCATACACCATTA	1187		
Qy	653	ACATTGATCTTACAGGAACACTATCCAAAGGTGATCTCCAAAGAGCTTTGATGATGATA	712		
Db	1186	ACATCGACATCTCCAGAGAACACTATATAGTAAAGTTCATCACCAAGAGAACCTTACGACGACA	1127		
Qy	713	GATTGAAAACATTCAAAACCGTTCAAGGATCTCGAATTAAGGCAATGACAGGTGGTATTC	772		
Db	1126	GGTTACAGACCATCAGAATGTCCAGAATCTCGAATTTGGAATTAAGGCTGTACCGGTGGTATTT	1067		
Qy	773	TTGGTCTTGGTGAGACCCAAAGAGGACCGTGTATCTTTCTCTACACCTTGGCCACCAATGG	832		
Db	1066	TGGGTCTCGGTGAAGCGAAGACCAATATAGGATTCATCTACACATTTCCATATGT	1007		
Qy	833	ATCAGCATCCAGATCTCTTCCAAATCAACAGACTGGTCCCAATCAAGGGCAACGCAATGT	892		
Db	1006	CTCCTCATCTCGTGTGCTCCCAATTAATAGACTAGTTGCTATCAAAAGGACCTCCCAATGG	947		
Qy	893	ATGAAGAAGTT-----AAGAAACAGCAAGTTGAAGTTGATGAGATTTGTCAGAACCA	943		
Db	946	CTGAGGAACCTTCCCGATCCAAAGAGTAAAAGTTGCAATTCGACGAATTTTGAGAACCA	887		
Qy	944	TTGCTACTGCAAGATTGGTCAATGCAACGCTCTATTATCAGATTGGCTGCAAGGAGATATA	1003		
Db	886	TTGCCACAGCGAATAATGTTATGCCAAGGCAATTAAGACTTTGCCGCTGGTCGTTATA	827		
Qy	1004	CAATGAAGAGGCGAAGACAGGTGATGTCTTCATGGCTGGTGTGATGCAATCTTCACAG	1063		
Db	826	CAATGAAGAAAACAGACAAATTTGTCTGTTTCATGGCAGAGTTGTAAACAGTATCTTCACCG	767		
Qy	1064	GTAAAGAAATGCTCACAAACAAATGTGTAACCGCTGGGATGAGGATTAAGCCATCTTCGCTA	1123		
Db	766	GTAAAGAAATGCTCACAGCAATGTGTAACCGTTGGGACGAGAACAGGCAATGTTGGCTA	707		
Qy	1124	AATGGGTCTGAAACCAATCGAGGTTTCAAAATACAA	1160		

593 TGGTGTAAAGACCAAGCACTGAATTTGAAAAGTGTGGGTGACGCGGTACCAACATA 652
 Db |||||
 16183 TGGTGTAAAGACCAAGCACTGAATTTGAAAAGTGTGGGTGACGCGGTACCAACATA 16242
 Db |||||
 593 ACATTGATCTTCAAGGAAACACTATCCAAAGGTGATCTCCACAAAGAGCTTTGATGATA 712
 Qy |||||
 16243 ACATCGACACTTCAGAGAAACACTATAGTAGGTGATCATCCACGAGAACTTACGACGACA 16302
 Db |||||
 713 GATTGAAACATTCAAAAAGTTCGAAGGATCTGGAATTAAGGCAATGACAGGTTGATTC 772
 Qy |||||
 16303 GGTTCAGACCATCAAGAAATGTCGAAGATCTGGAATTAAGGCAATGACAGGTTGATTC 16362
 Db |||||
 773 TTGGTCTTGGTGAGACCAAGAGGACCGTGTATCTTTCTCTACACCTTGCCCAATGG 832
 Qy |||||
 16363 TGGGTCTCGGTGAAGCGAAGACGACCATATAGGATTCATCTACACATTATCCAAATGT 16422
 Db |||||
 833 ATCAGCATCCAGAGTCTTCCATCAACAGACTGGTCCCAATCAAGGGCAGCGCAATGT 892
 Qy |||||
 16423 CTCCTCATCTCTGAGTCCCTACCAATTAATAGACTAGTTGCTATCAAGGGGACTCCAAATGG 16482
 Db |||||
 893 ATGAGAGATT-----AGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCA 943
 Qy |||||
 16483 CTGAGGAACTTGGCCATCAAGAGTAGTAAAGTTGCAATTCGACGAATTTTGAGAACCA 16542
 Db |||||
 944 TTGCTACTGCAAGATTGGTCAACAGCTCTATTTATCAGATTGGCTGCGAGGAAGATATA 1003
 Qy |||||
 16543 TTGCCACAGCGAATAGTTATGCCAAGGCCATTAATAGACTTGGCGCTGGTGTATA 16602
 Db |||||
 1004 CAATGAAAGGCGCAACAGGTGATGCTTCATGGCTGGTGTGAATGCCATCTTCACAG 1063
 Qy |||||
 16603 CAATGAAAGAACAGAGCAATTTGCTCTTTTCATGGCAGGTTGTAACAGTATCTTCACCG 16662
 Db |||||
 1064 GTAGAAATCTCACAACATGTGTACCGCTGGATGAGGATAAGCCATGTTGGCTA 1123
 Qy |||||
 16663 GTAGAAATCTCACAACATGTGTACCGCTGGATGAGGATAAGCCATGTTGGCTA 16722
 Db |||||
 1124 AATGGGTCTGAAACCAATGGAGAGTTTCAAAATACAA 1160
 Qy |||||
 16723 AATGGGATTGCAACCTATGAGGCAATTAAGTAGCA 16759
 Db |||||

RESULT 4

CR382121_05

WPCOMMENT

Sequence split into 11 fragments LOCUS CR382121 Accession CR382121

Fragment Name	Begin	End
CR382121_00	1	110000
CR382121_01	100001	210000
CR382121_02	200001	310000
CR382121_03	300001	410000
CR382121_04	400001	510000
CR382121_05	500001	610000
CR382121_06	600001	710000
CR382121_07	700001	810000
CR382121_08	800001	910000
CR382121_09	900001	1010000
CR382121_10	1000001	1062590

Continuation (6 of 11) of CR382121 from base 500001 (CR382121 Kluyveromyces lactis strain

Query Match 40.1%; Score 475.8; DB 8; Length 110000;
 Best Local Similarity 67.8%; Pred. No. 7.4e-117;
 Matches 685; Conservative 0; Mismatches 317; Indels 9; Gaps 1;
 175 TATGATTGTGAGTGAAGACAGCAGTCAACACCTGGACCAAGAGAAATTAAGCTATA 234
 Qy |||||
 46048 TATGACTATCTTTGAACGAACACAGCATGTGTGGACCAAGAACAAATTAAGCGAGATT 46107
 Db |||||
 235 TATGACACACCATCATGACTTGTGATGATGCTCAGGTGCAACAGAGGTTTCCAA 294
 Qy |||||
 46108 TATCACGCCATTGATGAAATGATGATCAGGCAATTAACATAGGAATGGCAT 46167
 Db |||||
 295 CAACCTTCAGAGGTTCAATTGTCACCTCTTATGAATATCAAACTGGTGGTTGACCGAG 354
 Qy |||||

RESULT 5

SCB102GNA

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

SCB102GNA 1847 bp DNA linear PLN 07-APR-1994
 S.cerevisiae gene for biotin synthetase.
 X72701
 X72701.1 GI:288356
 B102 gene; biotin synthetase.
 Saccharomyces cerevisiae (baker's yeast)
 Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;
 1 (bases 1 to 1847)
 Zhang, S., Sanyal, I., Bulboacă, G.H., Rich, A. and Flint, D.H.
 The gene for biotin synthase from *Saccharomyces cerevisiae*.

46168 GATCTCTACCAAGGTGCAACTGTGTACATTGATGAACATCAAGGAGGTGGATGTTTCAGAA 46227
 Db |||||
 355 GACTGTAAAGTACTGTGTCCCAATCAACAGCGTTAACAACACTGTGTCAAGGGTGAAGAATC 414
 Qy |||||
 46228 GATTGTAAAGTATTGTGGCGCAATCTTCAAGTTACGATACAGGTCTTAAAGCAGAGAATG 46287
 Db |||||
 415 ATCCAAGTTGATGAGGTGATGAGCTGCAAGAGGCAAGGCCAATGGAATCTTACAGG 474
 Qy |||||
 46288 ATATCTGCGAAGAAAGTGAATCAAGGAGGCAAAATTTGCCAAAGAAAATGGGTCCACGAGA 46347
 Db |||||
 475 TTCTGTATGGTCTGCTTCGAGAGAGATCAAGAGGTAGAAGTCAAACTTTGAAGAAAATC 534
 Qy |||||
 46348 TTTTGTCTTGGTCTGCTTCGAGAGAAATCGGTGTAGAANAATCTGCAATTTGAAGAAGATT 46407
 Db |||||
 535 AAAGAGATGATCACTGCTGTCCATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAATG 594
 Qy |||||
 46408 GGAGAATGATTGGAAAGATTACAGAGATGGGATGGAATCTTGGGTGACATTGGGAATG 46467
 Db |||||
 595 GTTGATAAAGACCAAGCCACTGAATTTGAAAAGTGTGGGTGTAGGGGTACGGGTACACCATTAAC 654
 Qy |||||
 46468 GTGGATGAAACTCAAGCGAAACAGTTTGAAGGACGCTGGTTTGACTGTCTTAAACCATTAAC 46527
 Db |||||
 555 ATTGATACCTTACAGGAGACACTATCCAAAGGTGATCTCCACAGAAAGCTTTGATGATAGA 714
 Qy |||||
 46528 ATCGATACCTCGAGGGGAACATTATCCCAAGGTGATCTCTACAGATCATATGACGATAGA 46587
 Db |||||
 715 TTGAAAACATTTCAAAAACGTTTCAAGGATCTGGAATTAAGGATGCAAGGTGGTGTATCTTT 774
 Qy |||||
 46588 TTGCAAAACCAATTAANAATGTGCAAGATGCAAGTATTAAAGCTTGCACAGGTGGATCTTG 46647
 Db |||||
 775 GGTCTTGGTGAGACCCCAAGAGGACCGGTATCTTTCTCTACCTTTGGCCACAAATGGAT 834
 Qy |||||
 46648 GGTCTTGGTGAAAACCTGAAGATGATCACATTTGGGTTCGTTTCCACACTTTTGGCAAAATATGAT 46707
 Db |||||
 835 CAGCATCCAGAGTCTCTTCAATCAACAGACTGGTCCCAATCAAGGCAAGCCCAATGTAT 894
 Qy |||||
 46708 CCACATCCGAATCTTTTACCTATTAATAGATAGTTTCCATTTAAAGGTACACCAATGAT 46767
 Db |||||
 895 GAAGAAGTTAAGAA-----CAAGCAAGTTGAAGTGTGATGATGATGATGATGATGATGAT 945
 Qy |||||
 46768 GAGGAATTAAGAAGGTGACAGCATGAAGAAATGGAATTCGATGAGATCTTAAAGAACCGTG 46827
 Db |||||
 946 GCTACTGCAAGATTGGTGTATGCGCAACGCTCTATTATCAGATTGGCTGAGGAGATATACA 1005
 Qy |||||
 46828 GCATAGCCAGAAATTTGAATGCCAACAGCGATTATCAGATTAGCTGCGAGTTCGTACACG 46887
 Db |||||
 1006 ATGAAAGAGGACAGAACAGGTGATGCTTTCATGGCTGTTGTAATGCCATCTTTCACAGGT 1065
 Qy |||||
 46888 ATGAAGGAAAACCTGAACAAATCTTTATGCTTTCATGGCAGGTTGTAATGCCATCTTTCGCGT 46947
 Db |||||
 1066 AAGAAAATGCTCAACAATGTGTAAACGCTGGGATGAGGATGAAGCCATGTTGGCTTAA 1125
 Qy |||||
 46948 AAAAAGATGTTAAACAATGTTGTACAGGCTGGGATGAGGATGAAGCAAGCAATGCTTTCGAA 47007
 Db |||||
 1126 TGGGCTCTGAAACCAATGAGAGATTTCAAATACAAACCAAGGGAGGTTGCA 1176
 Qy |||||
 47008 TGGGCTCTACCAACCAATGGAATCGTTTTCATATATCCGCTAATAAGGCTCCA 47058
 Db |||||

cloning, sequencing, and complementation of *Escherichia coli* strains lacking biotin synthase
Arch. Biochem. Biophys. 309 (1), 29-35 (1994)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

ORIGIN

Query Match

Best Local Similarity

Matches 694; Conservative

0; Mismatches 294; Indels 14; Gaps 3;

173

215

233

275

291

335

351

395

411

455

468

515

528

575

588

635

648

Query Match

Best Local Similarity

Matches 694; Conservative

0; Mismatches 294; Indels 14; Gaps 3;

173

215

233

275

291

335

351

395

411

455

468

515

528

575

588

635

648

Query Match

Best Local Similarity

Matches 694; Conservative

0; Mismatches 294; Indels 14; Gaps 3;

173

215

233

275

291

335

351

395

411

455

468

515

528

575

588

1. 1847
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="20B-12"
/db_xref="taxon:4932"
/chromosome="VII"
159.1226
/gene="BIO2"
159.1226
/gene="BIO2"
/codon_start=1
/product="biotin synthetase"
/protein_id="CAA51253.1"
/db_xref="GI:288357"
/db_xref="GOA:P32451"
/db_xref="UniProt/Swiss-Prot:P32451"
/translation="MPQLNRLHPOKLVPGSGSTICIVFRLTKSFVDKIAIKRNLSTPT
ARTVSCSHNCNRKWDPTKVLCTLMNKGSGCEDCKYCAQSRNDTGLKAEKMWK
VDEVTKRGRCGRKSTFRCGLAWRDMKGRKSAKRIQEWTKVNDMGLCTVTLG
MVDQAKQLKADGATATNHNIDTSREHSKYVITRTRVDDRLQITXNQESGIRKATG
GVLGSESDHIGFIYLSNPSHPESLPINRLVAIKGTMPAEELADPKKKLQFDE
ILRTATARIIVMPKAIIRLAAGRYTMKETBQFVCPMAGCNSIFTGKMLTIYNGWDE
DKAMLAKWGLQPMFAFKYDRS"

ORIGIN

Query Match 39.9%; Score 473.6; DB 8; Length 1847;

Best Local Similarity 69.3%; Pred. No. 3.2e-116;

Matches 694; Conservative 0; Mismatches 294; Indels 14; Gaps 3;

Qy 173 AATATGCAATGTCAGTGAAGCACCAGTCAACACCTGGACCAAGAAAGAAATTAAGCTA 232
Db |||||
Qy 215 AATATGCAATGTCAGTGAAGCACCAGTCAACACCTGGACCAAGAAATTAAGCTA 274
Db |||||
Qy 233 TATATGACACACCACTCATGCACTTGATGCACTATGCTCAGGTGCAA--CACAGAAAGTT 290
Db |||||
Qy 275 TTTATATACCCACTGCTCGAATTAATCTCATGCACTGCAATTCAGTCAAGTCAAGAAAGTG 334
Db |||||
Qy 291 CCAACAACTTCAGAGGTTCAATTTGTCATCTTATGAATATCAAACTGGTGGTTGTAC 350
Db |||||
Qy 335 GCACGATCCAAACCAAGTGAATTTGTCATGATGAATCAAACTGGTGGTTGTTC 394
Db |||||
Qy 351 CGAGGACGTGAAGTCTGTCCTGATCAAGGTTACACACTGCTGTCAGGTCAGGTCGAAG 410
Db |||||
Qy 395 TGAGGACGTGAAGTCTGTCCTGATCAAGGTTACACACTGCTGTCAGGTCAGGTCGAAG 454
Db |||||
Qy 411 AATCATCAAGTTGATGAGTGAATGA--AGCTGCAAGAGGCAAGGCAATGGATC 467
Db |||||
Qy 455 AATGTTAAAGTGAAGTGAATGAAGAGGCAAGAGGCTCAAAAGCAAGGATC 514
Db |||||
Qy 468 TACAAGTTCTGATGAGTGTGCTGTTGAGAGAGATGAAGGTGAAGTCAAACTTGA 527
Db |||||
Qy 515 TACTAGATTCCTGCTAGGTCGTGATGAGAGACATGAAGGTCTGAATCAGCCATGA 574
Db |||||
Qy 528 GAAATCAAGAGATGATCACTGCTGCTCCATGACCTTGAATGAGAGTGTGTCAACCT 587
Db |||||
Qy 575 AAGAAATTCAGAAATGGTGAACAAAGTGAATGATGGGGCTAGAAACGTGTGTACTTT 634
Db |||||
Qy 588 GGGAAATGGTTCATGAAGACCAAGCACTGAATTTGAAGAGTGTGGTGGAGCGGTACAA 647
Db |||||
Qy 635 AGGTATGGTTCATGAAGTCAAGCAAGCAATGAAGATGAGGTTGATGATGATCA 694
Db |||||
Qy 648 CCATAACATTCATCTTACAAGGAACACTATCCAAAGGTGATCTCCAAAGAGCTTTGA 707
Db |||||

RESULT 6

CR382130_18

WPCOMMENT

Sequence split into 37 fragments LOCUS CR382130 Accession CR382130

Fragment Name	Begin	End
CR382130_00	1	110000
CR382130_01	100001	210000
CR382130_02	200001	310000
CR382130_03	300001	410000
CR382130_04	400001	510000
CR382130_05	500001	610000
CR382130_06	600001	710000
CR382130_07	700001	810000
CR382130_08	800001	910000
CR382130_09	900001	1010000
CR382130_10	1000001	1110000
CR382130_11	1100001	1210000
CR382130_12	1200001	1310000
CR382130_13	1300001	1410000
CR382130_14	1400001	1510000
CR382130_15	1500001	1610000
CR382130_16	1600001	1710000
CR382130_17	1700001	1810000
CR382130_18	1800001	1910000
CR382130_19	1900001	2010000
CR382130_20	2000001	2110000
CR382130_21	2100001	2210000
CR382130_22	2200001	2310000
CR382130_23	2300001	2410000
CR382130_24	2400001	2510000
CR382130_25	2500001	2610000
CR382130_26	2600001	2710000
CR382130_27	2700001	2810000
CR382130_28	2800001	2910000
CR382130_29	2900001	3010000
CR382130_30	3000001	3110000
CR382130_31	3100001	3210000

Db 30513 AGATCATATGCGTTTCTTATACACCTTATCAACCATGGAAAAAGCATCCGGAATCTTTGCC 30454
 Qy 855 AATCAACAGACTGGTCCCAATCAAGGCGACGCAATGTATGAAGATTGAAG----- 907
 Db 30453 TATTAAACCGTTAGTGGCCATTAAGAAAGAACTCCCTGGAAGAAAGAGATATCCAGTTTACC 30394
 Qy 908 -----ACAAGCAAGTTGAAGTTGATGAGATTGTGAGAACCTTTGTCAGAACCTTTGTCAGAG 956
 Db 30393 TCGTGAATCAGGCGACAGCGTTTAAATTTGAACACATTTTAAACCAATTTGCTTACTGTAG 30334
 Qy 957 ATTGTCATGCCAAGCTTATTATCAGATTGCTGCTGAGGAAGATATACATGAAGAGGC 1016
 Db 30333 ATTAATTATGCCAGAGTCAATCATTAGGCTCGCTGCTGAGCGTTATACCATGAAGAAACA 30274
 Qy 1017 AGAACAGGTGATGCTTCATCGCTGGTGTGTAATGCCATCTTCAAGAGTGAAGAAATGCT 1076
 Db 30273 TGAGCAATCTTATGTTTTCATGAGTGGTGTGTAACGCTATCTTCAACGCTGAGAAATGTT 30214
 Qy 1077 CACAACAAATGTGTACGGCTGGATGAGGATAAAGCCATGTTGGCTAAATGGGGTCTGAA 1136
 Db 30213 AACCAATGTGTATGGTGGATGAAGACATTCAAATGTTTAAATAATGGGGTTGAG 30154
 Qy 1137 ACCAATGAGAGTTTCAATACAAACCAAGGAGGT 1172
 Db 30153 GCCTATGAAGAGCTTTAATAAAGCCATATGCCAGT 30118

RESULT 8
 AR549314
 LOCUS
 DEFINITION Sequence 4445 from patent US 6747137. linear PAT 08-OCT-2004
 ACCESSION AR549314
 VERSION AR549314.1 GI:53942489
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1263)
 AUTHORS Weinstock, K.G. and Bush, D.
 TITLE Nucleic acid sequences relating to *Candida albicans* for diagnostics and therapeutics
 JOURNAL Patent: US 6747137-A 4445 08-JUN-2004;
 FEATURES
 source
 1. .1263
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 34.9%; Score 414.4; DB 6; Length 1263;
 Best Local Similarity 66.8%; Pred. No. 2.8e-100;
 Matches 643; Conservative 0; Mismatches 301; Indels 18; Gaps 3;

Qy 207 CTGACCAAGAAAGAAATTAAGCTATATATGACACACCATCATGAGCTTGTGACACTA 266
 Db 297 CTGACCAAGAAATGAAATTTAGCAAAATTTTCAATGCTCCATGATGATTTGATTTCAA 356
 Qy 267 TGCTCAGGTGCAACACAGAAAGTTTCCACACCTTCAGAGGTTCAATGTGCACTTAT 326
 Db 357 CGCTCAATTGAAACATCGTGGAGTACCAAAACCTCGTGAAGTCCAAATATGATCTTGT 416
 Qy 327 GAATATCAAAACTGGTGGTGTCTACCGAGGACTGTAAGTACTGTGCCCAATCACAGCGTTA 386
 Db 417 TAACTTAAGCAGGCTGATGATGACGAAACTGTTCTTACTGTTCTCAATCATCGAAACA 476
 Qy 387 CAACACTGGTGTCAAGCTGAAAGAAATCATCAAGTTGATGAGGTGATGAAGCTGCAAA 446
 Db 477 TGATACTGGAGTTCAAGCTGAAAGAA--CTTGACTTGGATGCTGTGATGGACATTTGCTAG 533
 Qy 447 GGAGGCAAGGCCAATGATGATCTACAGGTTCTGATGGGTGCTGCTTTGGAGAGATGAA 506
 Db 534 AGATGCGAAAGCAGGAGGTGTAGAGTTTGGCTTGGTTCGCTTGGAGAGATATGCA 593

Qy 507 AGGTAGAAAGTCAAACTTTGAAGAAATAATCAAGAGATGATCACTGCTGTCATGAC---CT 563
 Db 594 AGGAAGAAAGTCTGCTCTAGACAAAATCAGTAAAGTTGTTTAAACAAAATCAACAGAGATT 653
 Qy 564 TGGAAATGAGAGTTGTGTCACCTTGGGAATGTTGATTAAGACCAAGCCACTGAATTGAA 623
 Db 654 GAATGTGGAACACTTGTGTACTCTCGGATGATCAACGAAAGAGCGGTGAAATTTTAAA 713
 Qy 624 AAGTCTGGGTGAGCGGTACAAACCAATAATGATGATCTTCAAGGAACACTATCCAAA 683
 Db 714 ACAGATGGTTTGAATGCTTACCAACCAACCAATGACACTTCAAGAGAACATTTATCCAA 773
 Qy 684 GGTGATCTCCAAAGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 743
 Db 774 CGTCTCACCACCGGTACATACGACGAAGATTTAGAAACCATTTAAACACGTCCAAAAAGC 833
 Qy 744 TGGATTAAGGCATGACAGGTGATTTCTTGGTCTTGGTGGAGACCCCAAGAGACCGTGT 803
 Db 834 TGGTATCAAGCTTGTACCGGTGATTTCTCGGATTTGGGAGAAACCGCACAGATCATGT 893
 Qy 804 ATCTTTCTCTACACCTTGGCCCAATGGATCGATCAGATCTCTTCCAATCAACAG 863
 Db 894 TTCGTTTTGTACACTTTTGTCCACATGATCCGACCCAGATCATTTGCCAATCAACAG 953
 Qy 864 ACTGTCCTCAATCAAGGCGACGCAATGTATGAAGAAAGT-----TAAGAAACA 911
 Db 954 ATTAGTTCCAATCCAGGTACTCCAATTTGTAAGTTTGAATAATGCCCAAGGAAG 1013
 Qy 912 GCAAGTTGAAGTTGATGAGATTGTGAGAACCAATTTGCTACTGCAAGATTGGTCAAGCAAC 971
 Db 1014 ACAATTTGAAATTTGATGCTATTTTGAAGAACTATTTGCTACTGATGATTTATGCTGTA 1073
 Qy 972 GTCTATTATCAGATTGCTGCGAAGATATACATGAAGAGGCGACAGAGGTGATGTG 1031
 Db 1074 ATCCATCATGATTGGCCCGGTGATACCATGAAGAAACACGAGAGTTCTTGTG 1133
 Qy 1032 CTTCATGGCTGGTTGTAATGCCATCTTCAAGGTGAAGAAATGCTCAACAAATGTGTAA 1091
 Db 1134 TTTCATGATGGGTCAATGCTATTTTACCGGTGAAGAGAAATTTGACTACAATGTGTAA 1193
 Qy 1092 CGGTGGATGAGATTAAGCCATGTTGGCTTAATGGGTCTGAACCAATGGAGAGTTT 1151
 Db 1194 TGGTTGGATGAAGATATACCATGTTTGAAGAAATGGGTTTGAACCAATGGAAAGTTT 1253
 Qy 1152 CA 1153
 Db 1254 CA 1255

RESULT 9
 SPCC320
 LOCUS
 DEFINITION S.pombe chromosome III cosmid c320. linear PLN 23-JUN-2003
 ACCESSION AL022245
 VERSION AL022245.2 GI:4584240
 KEYWORDS
 SOURCE Schizosaccharomyces pombe (fission yeast)
 ORGANISM Schizosaccharomyces pombe
 Schizosaccharomyces pombe
 Schizosaccharomycetales; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetales;
 Schizosaccharomycetes.
 1 (bases 1 to 40005)
 REFERENCE Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R.,
 Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D.,
 Bowman, S., Brooks, K., Brown, D., Brown, S., Chillingworth, T.,
 Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P.,
 Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D.,
 Hidalgo, J., Hodgson, G., Holtroyd, S., Hornsby, T., Howarth, S.,
 Huckle, E.J., Hunt, S., Jagsels, K., James, K., Jones, L., Jones, M.,
 Leather, S., McDonald, S., McLean, J., Mooney, P., Moulé, S.,
 Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neill, S.,
 Pearson, D., Quail, M.A., Rabinowitch, E., Rutherford, K., Rutter, S.,
 Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M.,

Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weijens, I., Vanstaels, E., Rieger, M., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritz, C., Holzer, E., Moesti, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dreano, S., Gloux, S., Lelaure, V., Mortier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Galliard, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrall, B.G. and Nurse, P.

The genome sequence of *Schizosaccharomyces pombe*
Nature 415 (6874), 871-880 (2002)

2 (bases 1 to 40005)

Wood, V., Rajandream, M.A., Barrell, B.G., Wedler, H., Wambutt, R. and Wedler, E.

Direct Submission
Submitted (24-MAR-1998) European Schizosaccharomyces genome sequencing project, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: pombe@sanger.ac.uk and AGOWA GmbH, Glienicke Weg 185, D-12489 Berlin, Germany

On Apr 14, 1999 this sequence version replaced gi:2995362.

Notes:
Details of *S. pombe* sequencing at the Sanger Institute are available on the World Wide Web.
(URL, <http://www.genedb.org/genedb/pombe/index.jsp>) (URL, http://www.sanger.ac.uk/Projects/S_pombe/)
CDS are numbered using the following system eg SPACSH10.0lc.SP (S. pombe), A (chromosome 1), cSH10 (cosmid name), .01 (first CDS), c (complementary strand). However, clones may have been reorientated since the original submission, therefore the complementary strand notation may be invalid for strand inference. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once.

Location/Qualifiers
1..40005
/organism="Schizosaccharomyces pombe"
/mol_type="genomic DNA"
/strain="972h"
/db_xref="taxon:4896"
/chromosome="III"
/map="III"
/clones="cosmid c320"
1..7687
/note="nominal overlap with cosmid SPCC330, EM:AL031603 S. pombe chromosome 3"
complement(1..32)
/note="rpl24-2 see c330 for full length CDS"
complement(709..1680)
/gene="SPCC320.14"
/note="synonym: SPCC330.15c"
complement(709..1680)
/gene="SPCC320.14"
/EC number="4.3.1.19"
/note="threonine dehydratase (predicted); involved in threonine catabolism; involved in branched chain amino acid biosynthesis; threonine dehydratase activity; localization unknown; similar to *S. cerevisiae* YKJ218C PATH:MAP00260;
aspect=P; term=threonine catabolism; GOid=GO:0006567; evidence=ISS; db_xref=GOC:unpublished, with=SPTR:P05792; date=20030609
aspect=P; term=branched chain family amino acid biosynthesis; GOid=GO:0009082; evidence=ISS; db_xref=GOC:unpublished, with=SPTR:; date=20030609
aspect=F; term=threonine dehydratase activity; GOid=GO:0004794; evidence=ISS; db_xref=GOC:unpublished;

with=SPTR:; date=20030609
aspect=C; term=cellular component unknown;
GOid=GO:0008372; date=20030609"
/codon_start=1
/protein_id="CAA18316.1"
/db_xref="GI:2995376"
/db_xref="GOA:O59791"
/translation="MSDNLVLTPTDDYDASERIKKFKANKTPVLTSTSTNKEFVAEVE
PCENFQMGAFKFGALNALSQLNEAQKAGVLTFFSGNHAQAIALSAKILGIPAKI
INPLDAPKAVATKYGYQVIMYDRYKDDREMAKEISEREGLTIIIPVDHPHVLG
QGTAAKELFEVGDALFVCGGGLSGSALAKHAFNPCEVGVVEAGNDGQGS
FRKSGIHDTPTKTADGAQTGGLGNYTFSIKERKVDILTIVSDEBILDLCKFYAARM
KIVVPTGCLSPAAARAMEKLNKRIIGIIISGNNVDIERYAHFLSQ"
complement(1123..1518)
/gene="SPCC320.14"
/note="Pfam match to entry S_T_dehydratase ,
beta-synthase"
2454..2525
/gene="SPCTRNAGLU.09"
2454..2525
/gene="SPCTRNAGLU.09"
/product="tRNA-Glu"
/note="tRNA Glutamic acid
tRNA Glu anticodon CTC, Cove score 74.15"
3149..4216
/gene="ark1"
/note="synonyms: sex1, SPCC320.13c, SPCC330.16"
3149..4216
/gene="ark1"
/note="serine/threonine protein kinase; aurora homolog;
involved in spindle checkpoint; involved in spindle
formation (required) (PMID 11792803); involved in sister
chromatid separation (required) (PMID 11950927); involved
in the microtubule attachment response (required) (PMID
12676091); interacts physically and/or genetically with
the inner centromere protein Plcp (PMID 11950927);
interacts physically and/or genetically with the inner
centromere protein Birp (PMID 11950927); interacts
physically and/or genetically with survivin (PMID
11950927); involved in protein amino acid phosphorylation
(SER10 of Htt3p) (PMID 11792803); similar to *S. cerevisiae*
YPL209C
08.01.3003 trimmed to 2nd met PMID 12676091"
/codon_start=1
/protein_id="CAD88264.1"
/db_xref="GI:29786364"
/translation="MSDSKLADSLNCLSVSTSTTANPGRQQLLRVLAVSNQRQNVVS
LANGENKRTSNKSTNSLRKTEEFVAGVSSAGPQWREFHGMFEGPLGKFGFR
VYLAKKFTGVALKTLKSELVQSKIEKQVRETEIQSNLRHKNILRYLGHFDEK
RIVYLLEPAGRGELYOHLRRARRFSEEVASKYIFOMANALSYLHKHVIHRDIKPNKI
LLIGDGEIKLSDPGMSVHAPSRRRTLCGTLDYLPPEVGEKHEKVDLWSLGLVLY
EFLVGAPPEDMSGHSATYKRIAKVDLKIPSPFPDPARDLISRLQLHNFKEKMSLSQV
MRHPWIKYKDSWTRKSSS"
3149..3151
/gene="ark1"
/note="initiator methionine, confirmed PMID 12676091"
3413..4168
/gene="ark1"
/note="Pfam match to entry pkinase PF00069, Eukaryotic
prot ein kinase domain"
3431..3502
/gene="ark1"
/note="PS00107 protein kinases ATP-binding region
signature"
3770..3808
/gene="ark1"
/note="PS00108 Serine/Threonine protein kinases
active-site signature"
complement(join(4707..4714,4781..5330))
/gene="SPCC320.12"
/note="synonym: SPCC330.17c"
complement(join(4707..4714,4781..5330))

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

FEATURES
source

misc_feature

misc_feature

gene

CDS

Query Match	28.9%	Score 343.2	DB 8	Length 40005
Best Local Similarity	60.7%	Pred. No. 3.7e-81		
Matches 580	Conservative 0	Mismatches 373	Indels 3	Gaps 1
Qy	202	AACACCTGGACCAAGAAGAAATTAAGCTATATATGACACACACACTCATGGACTTGATG	261	
Db	38504	ACAATTGGACTCGTGAAGAAATTCAAAAATTTACGACACACATTGATCGAATTTGATT	38563	
Qy	262	CACATGCTCAGGTGCACACAGAAGGTTCCACAACACTTTCAGAGGTTCAATTGTGCACCT	321	
Db	38564	TTTAGACCGCTCCCATACATAGAAATTTCCACGCCCAAGAAGTTCAACAATGCACCT	38622	
Qy	322	CTTATGAATATCAAACTCGTGGTTGTGACGAGGACTGTAAAGTACTGTGCGCCAAATCACAG	381	
Db	38624	TTACTTTCTATTAATAACAGGTGGATGTACAGAAGATTGCAAAATATTGTGCGCAATCTTCC	38683	
Qy	382	CGTTACAACACTCGTGTCAAGGCTGAAAGAAATCATCCAAGTTGATCAGGTGATTGAAGCT	441	
Db	38684	CGCTACAATACCGGTGTCAAGGCCACTTAAGCTCATGAANAATTGACGAAGTTTTCGAAAAAG	38743	
Qy	442	GCAAGGAGGCAAGGCGCAATGGATCTACAAGGTTCTGTATGGGTGTCTTTCGAGAGAG	501	
Db	38744	GCTAAAATTCGAAGGCTAAGGAGCCTAGATTTGTATGGGAAGTGCCTGGCGTGAT	38803	
Qy	502	ATGAAGGTGAGAAGTCAAACTTGAAGAAATCAAGAGATGATCATCTGCTGTCATGAC	561	
Db	38804	TTGAATGGTGCCAATCGCACCTTTTAAAAAATTTTGGAAATCATCAAGAGGTTTCGCAGC	38863	
Qy	562	CTTGGAAATGAGAGTTGTGTGTCACCTGGGAATGGTTGATAAGACCAAGCCACTGAATTG	621	
Db	38864	ATGGATATGGAAGTTTTCGTTAACCTTTGGTATGTTTAATGACAGCAGCCAGGAACTA	38923	
Qy	622	AAAAGTCTGGGTGACGGCGTCAACACCAATAACATTTGATATCTTCAAGGAACACTATCCA	681	
Db	38924	AAGGATGCAGGCTTGACGGCATACAATCATTAATCTTGATACCTCTCGCGAATATTACTCA	38983	
Qy	682	AAGGTGATCTCCAAGAAGCTTTGATGATAGATTGAAAACATTTCAAAAACGTTCAAGA	741	
Db	38984	AAGATTATCAGCACTAGGACATATGATGACCGTTAAATATCTATTGCAATCTTCGGA	39043	

FORMED
REFERENCE 2 (bases 1 to 40880)

AUTHORS Wood, V., Rajandream, M.A., Barrell, B.G., Murphy, L. and Harris, D.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1998) European Schizosaccharomyces Genome
 sequencing project, Sanger Institute, The Wellcome Trust Genome
 Campus, Hinxton, Cambridge CB10 1SA, E-mail: pombe@sanger.ac.uk

Notes:
 Details of S. pombe sequencing at the Sanger Institute are
 available on the world wide web.
 (URL, http://www.sanger.ac.uk/projects/s_pombe/)
 CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
 pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
 (complementary strand). However, clones may have been reorientated
 since the original submission, therefore the complementary strand
 notation may be invalid for strand inference. IMPORTANT: This
 sequence MAY NOT be the entire insert of the sequenced clone. It
 may be shorter because we only sequence overlapping sections once.
 Location/Qualifiers

FEATURES
 source
 1. .40880
 /organism="Schizosaccharomyces pombe"
 /mol_type="genomic DNA"
 /strain="972h-"
 /db_xref="taxon:4896"
 /chromosome="III"
 /map="III"
 /cosmid="c1235"
 1. .2935
 /note="nominal overlap with cosmid SPCC320, EM.AL022245 S.
 pombe chromosome 3"
 1. .576
 /gene="SPCC1235.01"
 /note="synonym: SPCC320.02c"
 <1. .576
 /gene="SPCC1235.01"
 /codon_start=1
 /protein_id="CAA21105.1"
 /db_xref="GI:3668148"
 /db_xref="UniProt/TREMBL:O59779"
 /translation="DLAKSLFONRNLTLKLAQGLVLTAAQSGGSLGSLAAVSGLLGS
 FMSSNNSSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNS
 NNNNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNSNNNS
 1374. .2465
 /gene="bio2"
 /note="synonyms: SPCC1235.02, SPCC320.01c"
 1374. .2465
 /gene="bio2"
 /EC_number="2.8.1.6"
 /codon_start=1
 /protein_id="CAA21106.1"
 /db_xref="GI:3668149"
 /db_xref="GOA:O59778"
 /db_xref="UniProt/Swiss-Prot:O59778"
 /translation="MFRITRQIRRSALSLVNNWTRBEIQIYDTPILDLIFRAA
 IAKRHPDPRKVOOCTLLSRTGGCTEDCKYCAOSSRYNTGVKATLMKIDLEVLAK
 IAKAGTRFCWGSAMDRLGRNTPFNILEITKEVSDMEVCVTLGMLNEQQAKEL
 KDAGLTAYNHLDSREYYSKIISTRYDRLANTIDLRKAGLVKSGGILGLGKXK
 DRVGLIISLTAMTPHPSVPFNILVPIPGTPGVDAVKERLPHPFLRSIATRICMPK
 TIIPRAGNRTCSSEQALAFAGANAVFTGERKMLTTPAYSWDSDSLFWNWGLEGNQ
 SFYGTSTREGDGFITLPKLERLAPFSL"
 1434. .2366
 /gene="bio2"
 /note="Match to Biotin synth , Score 756.53"
 3804. .5230
 /gene="SPCC1235.03"
 3804. .5003
 /gene="SPCC1235.03"
 /note="fmr domain protein (small muts related protein);
 DNA repair related; no apparent orthologs"
 /codon_start=1
 /protein_id="CAA21107.1"
 /db_xref="GI:3668150"
 /db_xref="GOA:O74840"

misc_feature
 3'UTR
 /db_xref="UniProt/TREMBL:O74840"
 /translation="MPSDLSAARVQLVQKCAQNDLVRSSDELNHRRAIQESQIETAA
 DALITSNIGHRQRKTKKATNSKPLSKFQNTSEVNEDEPILKPSLSVWNNRLL
 IETLITNIPSSQINHEFYKNSAWPIITRNLIRHQPISLNTNNELMKYQBEATOL
 AKDTGLSLKICTDLVICSNDYKNALWILCLIKETOHNMGNIKLSOTAKSNSTOTK
 TCLNDSSKLVEDDEALSADCNRAEYELRNMOYSNAKEYRSKSNHLFGGSAM
 YHAQGREYERKALRSLAMRSLAHSGLSHLDLHGATVREAKTIVRERVAWAKK
 ADTSPGRIHFPIVITGRGNSIGLEARLLPAIVRLQQDHRFDEAHGQITVYGINRH
 SKLSNSNA"
 5004. .5230
 /gene="SPCC1235.03"
 /note="mRNA from SP0171"
 complement(join(5485. .5736,5785. .6193,6247. .6383))
 /gene="SPCC1235.04c"
 complement(join(5485. .5736,5785. .6193,6247. .6383))
 /gene="SPCC1235.04c"
 /EC_number="2.7.7.2"
 /note="PAD synthetase (predicted); similar to S.
 cerevisiae PAD1"
 /codon_start=1
 /protein_id="CAA21108.1"
 /db_xref="GI:3668151"
 /db_xref="GOA:O74841"
 /db_xref="UniProt/Swiss-Prot:O74841"
 /translation="MDELERYQSIKNIETHSAPSEKLYVLQNLRLSISLRFIEYAPET
 YOPERLAMSFGNGKCLVLFLICLYLKEKYEQAKLSNIPFVFRPDEPPEMDD
 FYNKOSKYLRLNIKIISLPMKFAFCFLKEHKHQAILIGIRLDPHGLHRIAFVTD
 KGWPKMRITQIILDSRYNIWDLLETNTKYCSLYDRGYTSLGGVSDTSPNALKNPD
 GYSPAYLLSDGSLERAGNNTVPFSRTNSRFLYDSGASSYASVS"
 complement(5737. .5747)
 /gene="SPCC1235.04c"
 /note="ctaacattag, splice branch and acceptor"
 complement(5779. .5784)
 /gene="SPCC1235.04c"
 /note="gtaagt, splice donor sequence"
 complement(6194. .6210)
 /gene="SPCC1235.04c"
 /note="ctaacgtttattacaag, splice branch and acceptor"
 complement(6241. .6246)
 /gene="SPCC1235.04c"
 /note="gtagtc, splice donor sequence"
 complement(7185. .11294)
 /gene="SPCC1235.05c"
 complement(7185. .7439)
 /gene="SPCC1235.05c"
 /note="mRNA from AU013698"
 complement(7440. .11294)
 /gene="SPCC1235.05c"
 /note="DEAD/DEAH box helicase, SNF2 family; helicase
 C-terminal domain; involved in UV protection (predicted);
 similar to S. cerevisiae SNF2; paralog?"
 /codon_start=1
 /protein_id="CAA21109.1"
 /db_xref="GI:3668152"
 /db_xref="GOA:O74842"
 /db_xref="UniProt/TREMBL:O74842"
 /translation="MLPYNSLSDNGKFKSDENPQSEVYSGSQTLPSYSGNPOSY
 GTPVQOQSAMTGVNMGGMNTSENQFMNTDYQTSYSSYASTPMSNAYSRDAPAA
 INNFGYSYVGOSQPVPSYNPLPSYNTASLNAGIIPAAMPGPSYGTPTVPOGGY
 NAHSSPNYNGYPIGAVNTSAIPQAPPAQVNNVLPSVRSNRSASSTARSARS
 TORSSSNPNVPTTPTNNLTTPAPPVPELPTVTTSNALRSQVHRSVPQAPI
 HVQINTLAQKWDETAALSILSKYLSCLDGIPIQHKRFQSPVASNMPTGSSNRT
 VQSKRSIRDKYIQMNDSTOASLMPSTYRTKTSNASKLTTEDEFPYDESEPEALVH
 RDTLSALERTLVNFINSSTAKSLDSTASCPYSHKLLLEHFPQTLAEACIIKHDDVP
 SKPGRGRERKNPMGOKIVNACMETMEGYAIDNLIKACEFLGNRIKSGMASWGIKL
 EMSNGELNIVMESVPTAEADNDPFPKTEQPKTLASDQSLKQVGVNNHLLYQ
 QKLSGLADAMEGLKTCQVAFALLLEQGHGPHLVVPSSFTLENLWRELARFCPSL
 GYLKNNMSERYKHLNLANFNRLLTGTLQNNLKLVSLLAFILPNPFSDMDOLD
 VIFKAPKADADIQALLSKORISRAKTMWTFLVRKRNQVNLDPKTKTQIIEHCKL
 SENOLEIYNYAALOKNQLRRDDKKNKESKNDSDGSLGAGHVLMLRKAANHAL
 LPRKVDDEKLQMAKDINQEQYKNANEQYLYEDMEVMSDFELHRLCHSFPTLQSYT
 LKDDPMDWGGKIRVLKELLPKMKEGSRILLFSQFTQMLDILLEQVLDLTKI SYVRLDG

STOVEVRDIIIDQFHEEDVTVFLSTAGGPGINLACANVILVDCSYNPFDDLOAE
DRAHVGQVRETVIRLIDNTIEYIKLANIKLADMSLSSDGGREERGERLUVQD
MLDENGNNTKPEITIGNESDGEFKVSSNNKQDABETNTGVPLSGQPNSEKVD
LADGEKANIETKMSSETVEGDNKELRETKGENTVDNNAVPSKSESTERNESVL
SGHLDLDEASPVWSTIEKTKGDSVTEBQOSANIDGQLEKPEIEESKPDVLNOVS
LSIEBKPNKESVEDNNAKD"
complement (8184..8435)
/gene="SPCC1235.05c"
/note="Match to helicase_C , Score 83.11"
complement (8721..9638)
/gene="SPCC1235.05c"
/note="Match to SNF2_N , Score 344.47"

misc_feature

misc_feature

Query Match 28.9%; Score 343.2; DB 8; Length 40880;

Best Local Similarity 60.7%; Pred. No. 3.7e-81;

Matches 580; Conservative 0; Mismatches 373; Indels 3; Gaps 1;

Qy	202	AACACCTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGACCTGATG	261
Db	1434	AACAATGGACTCGTGAAGAAATTCAAAATAATTACGACACACCACTGATGATTTGAT	1493
Qy	262	CACATGCTCAGGTGCAACACAGAGAGGTTCCAAACCACTTCAGAGGTTCAATTTGCACT	321
Db	1494	TTAGAGCGCTTCATACATAGAAATTCACGACCCCAAGAAAGTTCAACAATGCAT	1553
Qy	322	CTTATGAATACAAACTGGTGTGTACCGAGGACTGTAAGTACTGTGCCCCAATCACAG	381
Db	1554	TTACTTTCTATTAAACACAGGTGGATGTACAGAAGATTGCMAATATTGTGCGCAATCTCC	1613
Qy	382	CGTTACACACTGGTGTCAAGCTCAAGATCATCCAAAGTGTGAGAGTGTGAAGCT	441
Db	1614	CGCTACAATACGGTGTCAAGGCCACTAAGCTCAAGAAATTTGACGAAGTTTGGAAAG	1673
Qy	442	GCAAAGGAGGCAAGCCCAATGGATCTACAAGGTTCTGTATGGGTGCTGTTGAGAGAG	501
Db	1674	GCTAAATTCAGAGCTTAAGAGAGCACTAGATTTTGTATGGGAAGTGCCTGGGCTAT	1733
Qy	502	ATGAAGGTGAAGTCAAACTTGGAAGAAATCAAGAGAGATGATCACTGCTGTCTCATGAC	561
Db	1734	TTGAATGTCGCAATCGCACCTTTTAAATAATTTTGGAAATCATCAAGAGAGGTTGCGACG	1793
Qy	562	CTTGAATGAGAGTGTGTACCTCGGATGTTGATTAAGACCAAGCCATGAATG	621
Db	1794	ATGGATATGGAAGTTTGGCTAACCTCTGTATTTAAATGAACAGCAAGCCAAAGAACTA	1853
Qy	622	AAAAGTGTGGTGTGACGGGTACAAACATACATTTGATCTTCAAGAAACACTATCCA	681
Db	1854	AGGATGAGGCTTGACGGCATACATCATATCTTGAATCTCTCGGCAATATTACTCA	1913
Qy	682	AAGGTGATCTCCACAAGAAGCTTTGATGATAGATTGAAAAATTCAAAAAGCTTCAAGGA	741
Db	1914	AAGATTATCAGCACTAGGACATATGATGAACGGTTAAATACTATTGACAAATCTTCGAAA	1973
Qy	742	TCTGATTAAGGCATGACAGGTGGTATCTTGTCTTGTGTGACCAAGAGAGACCGT	801
Db	1974	GCTGGTTGAAGGTTTGTAGCGGTGGTATTTGGTCTTAGGGGAGAAAAGCATGATCGT	2033
Qy	802	GTATCTTCTCTACACCTTGCCCAATGATGATCAGATCCAGATCCAGATCTCTTCCATCAAC	861
Db	2034	GTAGATTGATTCATCTGTTAGCAACTATGCCAACATCTCTGATCTGTACCTTTAAT	2093
Qy	862	AGACTGTGTCCTCAATCAAGGCGACGCAATGTATGAAGAAGTTTAAGAAACAAGATTTGAA	921
Db	2094	CTTTTGTGTCGGATTCCTGGCACCCCTGT---AGGAGACGCTGTAAAAAGAGAGACTTCCG	2150
Qy	922	GTTGATGATGTGTCAGAACATGCTACTGCAAGATTGGTCATGCCAACGCTCTATTATC	981
Db	2151	ATTCTATCTTTCTTCTGCTCCTCACTGCAACGCTCGTATTTTGCATGCCAAAGACTATCAT	2210
Qy	982	AGATTGGCTGAGGAAGATATACAATGAAGAGGACAGAAAGAGTGTGCTGCTCATGCT	1041
Db	2211	CGTTTGACGCTGACGTAACGCTGCTCCGAGTCCGAGCAAGCACTTGTCTTTCATGCT	2270

Qy	1042	GGTTGTAATGCCATCTTACAGGTAAGAAATGCTCAACAATGTGTACGGCTGGAT	1101
Db	2271	GGTGAATCAGATTTTACCGGTGAAATAATGCTTACTCTCTGCTGTTCTTTGGGAC	2330
Qy	1102	GAGGATAAGCCATGTTGGCTAAATGGGCTCTGAAACCAATGAGAGATTTCAATA	1157
Db	2331	TCAGACAGTACGTTTTTACAAATTTGGGTTGGAGGAATGCAAGGTTTGAATA	2386

RESULT 11

LOCUS

CNS06K9X 847 bp DNA linear STS 10-JAN-2001
T3 end of clone AT0AA003D06 of library AT0AA from strain CBS 4311
of Saccharomyces servazzii, sequence tagged site.

ACCESSION

VERSION

AL402619

KEYWORDS

AL402619.1 GI:12161819

SOURCE

Saccharomyces servazzii

ORGANISM

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE

1 (bases 1 to 847)

AUTHORS

Souciet J L, Aigle M, Artiguenave F, Blandin G,
Bolotin-Fukuhara M, Bon E, Brottier P, Casaregola S,
de-Montigny J, Dujon B, Durrens P, Lepingle A, Liorente B,
Malpertuy A, Neuvéglise C, Ozier-Kalogeropoulos O, Potier S,
Saurin W, Tekala F, Toffano-Nioche C, Wesolowski-Louvel M,
Wincker P, and Weissenbach J.

TITLE

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

JOURNAL

FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE

20584711

PUBMED

11152876

REFERENCE

2 (bases 1 to 847)

AUTHORS

Casaregola S, Lepingle A, Bon E, Neuvéglise C, Nguyen H,
Artiguenave F, Wincker P, and Gaillardin C.

TITLE

Genomic exploration of the hemiascomycetous yeasts: 7.

JOURNAL

FEBS Lett. 487 (1), 47-51 (2000)

MEDLINE

20584717

PUBMED

11152882

REFERENCE

3 (bases 1 to 847)

AUTHORS

Genoscope.

Direct Submission

Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]

TITLE

This STS is part of a random genomic sequencing program of thirteen

yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces

exiguus, Saccharomyces servazzii, Kluyveromyces thermotolerans, Kluyveromyces

lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia

angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,

Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of

the other extremity of this insert.

FEATURES

Location/Qualifiers

1..847

/organism="Saccharomyces servazzii"

/mol_type="genomic DNA"

/strain="CBS 4311"

/db_xref="taxon:27293"

/clone_lib="AT0AA003D06"

/clone="AT0AA003D06"

/note="end : T3"

complement (<8..>847)

/note="similar to Saccharomyces cerevisiae ORF YGR286c [

BI02 : biotin synthetase]"

/evidence=not_experimental

ORIGIN

Query Match 28.0%; Score 332.2; DB 11; Length 847;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 14:47:19 ; Search time 731 Seconds
(without alignments)
9620.592 Million cell updates/sec

Title: US-09-752-957F-1

Perfect score: 1188

Sequence: 1 atgcgtttattatgactgc.....agggtgcattcggtgcttga 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	9	ADA09588 Candida u
2	417.6	35.2	1125	12	ADP98553 C. albica
3	349.2	29.4	1659	6	ABA01978 Soybean b
4	349.2	29.4	1659	10	ADE76391 Ade76391 Soybean S
5	328.2	27.6	3083	2	AAT90969 Zuoatin ge
6	328.2	27.6	3083	2	AZ11218 Human zuo
7	328.2	27.6	3083	10	AAD59051 Yeast zuo
8	328.2	27.6	3083	12	ADO04472 Yeast Z-D
9	328.2	27.6	3083	13	ADS15330 Yeast DNA
10	311.8	26.2	1515	6	ABA01974 Maize bio
11	311.8	26.2	1515	10	ADE76383 Corn SID2
12	310.2	26.1	1313	3	AAC48072 Zea mays
13	310.2	26.1	1396	6	ABA01972 Barley bi
14	310.2	26.1	1396	6	ADE76379 Barley bi
15	310.2	26.1	1439	6	ABA01975 Maize bio
16	310.2	26.1	1439	10	ADE76385 Corn SID2
17	306.6	25.8	1152	6	ABA01967 Maize bio
18	306.6	25.8	1152	10	ADE76369 Corn biot
19	304	25.6	1137	6	ABZ12774 Arabidops
20	304	25.6	1137	10	ABV76106 Arabidops

21	304	25.6	1137	12	ADN74336
22	304	25.6	1340	3	AAC48445
23	304	25.6	1351	2	AAT34580
24	304	25.6	1351	2	AAX01311 A. thalia
25	300.8	25.3	1342	3	AAC33135
26	295.6	24.9	1477	6	ABA01976
27	295.6	24.9	1477	10	ADE76387
28	294.4	24.8	1340	6	ABA01969
29	294.4	24.8	1340	10	ADE76373
30	292.8	24.6	1526	6	ABA01977
31	292.8	24.6	1526	10	ADE76389
32	292.6	24.6	1041	1	AAN91327
33	285.6	24.0	1248	13	ADR85519
34	285.6	24.0	1248	13	ADR84932
35	285.6	24.0	7248	13	ADR84345
36	260.8	22.0	1467	6	ABA01973
37	260.8	22.0	1467	10	ADE76381
38	243.4	20.5	1032	6	ABA01979
39	243.4	20.5	1032	10	ADE76393
40	237.6	20.0	987	8	ACA20541
41	236	19.9	1017	9	ADA31411
42	228.6	19.2	957	8	ACA37222
43	226.4	19.1	1038	10	ACF68915
44	226.4	19.1	110000	10	ACF67367_17
45	226.4	19.1	210710	10	ACF65380

ALIGNMENTS

RESULT 1

ADA09588
ID ADA09588 standard; DNA; 1188 BP.

XX AC ADA09588;

XX DT 06-NOV-2003 (first entry)

XX DE Candida utilis biotin synthase gene.

XX KW Yeast; high biotin-productivity; edible yeast; biotin synthase gene;
XX RW feed additive; food additive; cosmetic; gene; ds.

XX OS Pichia jadinii.

XX PN US2003104584-A1.

XX PD 05-JUN-2003.

XX PF 02-JAN-2001; 2001US-00752957.

XX PR 07-OCT-2000; 2000TW-00120972.

XX PA (SHIU/) SHIUAN D.

XX PI Shiuian D;

XX DR WPI; 2003-567956/53.

XX PT New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.

XX PS Claim 3; Page 3; 20pp; English.

XX CC The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (Candida utilis) biotin synthase gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents C. utilis biotin synthase gene.

SQ	Sequence	1188 BP; 376 A; 229 C; 295 G; 288 T; 0 U; 0 Other;
	Query Match	100.0%; Score 1188; DB 9; Length 1188;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1188; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGTCGGTTTATATTGACTGTCTATTAGTCGTCGGAATTCCTCTTTCCACCTTCTAGAGTAGCT 60
Db	1	ATGTCGGTTTATATTGACTGTCTATTAGTCGTCGGAATTCCTCTTTCCACCTTCTAGAGTAGCT 60
Qy	61	TCTAGGCGCTACTCTGGCAACAGGTGCCTACTGCTCTCGGAGATCTTTGGAAAGATGCTGTC 120
Db	61	TCTAGGCGCTACTCTGGCAACAGGTGCCTACTGCTCTCGGAGATCTTTGGAAAGATGCTGTC 120
Qy	121	ACGGAACAAATGGAAAGATGGCTTTCACAGGAGAAGAGCCAAACCATTGGAAATATGCA 180
Db	121	ACGGAACAAATGGAAAGATGGCTTTCACAGGAGAAGAGCCAAACCATTGGAAATATGCA 180
Qy	181	TTGTCAGTGAAGACAACAGTCAACACCTGGACCAAAGAAGAAATTAAGCTATATATGAC 240
Db	181	TTGTCAGTGAAGACAACAGTCAACACCTGGACCAAAGAAGAAATTAAGCTATATATGAC 240
Qy	241	ACACCACTCATGGACTTGATGCACCTATGCTCAGGTGCAACACAGAAGSTTCCAACAACCT 300
Db	241	ACACCACTCATGGACTTGATGCACCTATGCTCAGGTGCAACACAGAAGSTTCCAACAACCT 300
Qy	301	TCAGAGGTTCAATTTGTGCACCTCTTATCAATATCAAAACTGGTGGTTGTACCGAGGACTGT 360
Db	301	TCAGAGGTTCAATTTGTGCACCTCTTATCAATATCAAAACTGGTGGTTGTACCGAGGACTGT 360
Qy	361	AAGTACTGTGCCAATCACAGCGTTTCAACACTGGTCAAGGCTGAAGAATCATCCAA 420
Db	361	AAGTACTGTGCCAATCACAGCGTTTCAACACTGGTCAAGGCTGAAGAATCATCCAA 420
Qy	421	GTTGATCAGGTGATTGAAGCTGCAAGGCGCAAGGCCAATGCACTCAAGGTTCTGT 480
Db	421	GTTGATCAGGTGATTGAAGCTGCAAGGCGCAAGGCCAATGCACTCAAGGTTCTGT 480
Qy	481	ATGGGTCTGCTCGGAGAGAGATCAAAAGGTAGAAGTCAAACTTGAAGAAAAATCAAAAG 540
Db	481	ATGGGTCTGCTCGGAGAGAGATCAAAAGGTAGAAGTCAAACTTGAAGAAAAATCAAAAG 540
Qy	541	ATGATCACTGCTGTCATGACCTTTGGAAATGGAGAGTTGTGTCAACCCTGGGAATGGTTGAT 600
Db	541	ATGATCACTGCTGTCATGACCTTTGGAAATGGAGAGTTGTGTCAACCCTGGGAATGGTTGAT 600
Qy	601	AAAGACCAAGCCA CTGAATTTGAAAAGTCTCGGGTTGACGGCGTACACCAATAACATTGAT 660
Db	601	AAAGACCAAGCCA CTGAATTTGAAAAGTCTCGGGTTGACGGCGTACACCAATAACATTGAT 660
Qy	661	ACTTACAAGGAACACTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTCAAA 720
Db	661	ACTTACAAGGAACACTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTCAAA 720
Qy	721	ACATTTCAAAAACGTTCAAGGATCTGGATTTAAAGGCATGCAAGGTGGTATTCTTGGTCTT 780
Db	721	ACATTTCAAAAACGTTCAAGGATCTGGATTTAAAGGCATGCAAGGTGGTATTCTTGGTCTT 780
Qy	781	GGTGAGACCCAAAGAGACCGTGATCTTTTCTCTACACCTTTGGGCCAACATGGATCAGCAT 840
Db	781	GGTGAGACCCAAAGAGACCGTGATCTTTTCTCTACACCTTTGGGCCAACATGGATCAGCAT 840
Qy	841	CCGAGGTCTCTTCCAAATCAACAGACTGTGTCCTCAATCAAGGSCACGCAATGTATGAAGAA 900
Db	841	CCGAGGTCTCTTCCAAATCAACAGACTGTGTCCTCAATCAAGGSCACGCAATGTATGAAGAA 900
Qy	901	GTTTAAGAACCAAGCAAGTTGAAGTTGATGAGATTGTCRGAACCAATTGCTACTGCAAGATTG 960
Db	901	GTTTAAGAACCAAGCAAGTTGAAGTTGATGAGATTGTCRGAACCAATTGCTACTGCAAGATTG 960
Qy	961	GTCATGCCAAGCTCTATTATCAGATTGGCTCGAGGAAGATATACAATGAAGAGCGAGAA 1020
Db	961	GTCATGCCAAGCTCTATTATCAGATTGGCTCGAGGAAGATATACAATGAAGAGCGAGAA 1020

specification (ADP98516-ADP98825); identifying a gene that is essential to the survival or growth of a fungus, that contributes to the virulence and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus, or a therapeutic agent for treatment of a mammalian disease; correlating changes in the levels of proteins or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of *Candida albicans*, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP99135; producing a polypeptide; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825; eliciting an immune response in an animal; a strain of *Candida albicans*, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP99135, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of *Candida albicans*; inhibiting growth or proliferation of *Candida albicans* cells; manufacturing an antimycotic compound; treating an infection of a subject by *Candida albicans*; preventing or containing contamination of an object by *Candida albicans*, or for preventing or inhibiting formation on a surface of a biofilm comprising *Candida albicans*; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or at least one amino acid sequence selected from ADP9826-ADP99135; a method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one protein comprises an amino acid sequence or a portion of an amino acid sequence selected from ADP98516-ADP98825. The novel methods and compositions have fungicide activity. The compositions may be used in gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated with *Candida albicans*. These may also be used for constructing strains useful for identification and validation of gene products as effective targets for therapeutic intervention, for identifying and validating gene products as effective targets for therapeutic intervention, and for collecting identified essential genes. This polynucleotide sequence represents a *Candida albicans* fungal specific gene of the invention.

NOTE: This sequence was downloaded from an electronic sequence listing provided on the WIPO website.

XX SQ Sequence 1125 BP; 362 A; 218 C; 242 G; 303 T; 0 U; 0 Other;

Query Match 35.2%; Score 417.6; DB 12; Length 1125;
Best Local Similarity 67.0%; Pred. No. 6.7e-113;
Matches 64S; Conservative 0; Mismatches 299; Indels 18; Gaps 3;

QY 207 CTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGGACTTGTGCACTA 266
DB 159 CTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGGACTTGTGCACTA 218
QY 267 TGCTCAGGTGCAACAGAGGTTCCAAACCTTCAGAGGTTCAATGTGCACTTAT 326
DB 219 CGCTCAATTGAACATCGTGAGTACCAAAAACCTGCTGAAGTCCCAATTATGACTCTGT 278
QY 327 GAATATCAAACTCGTGGTGTACCGAGGAGCTGTAAGTACTGTCCCAATCACAGCGTTA 386
DB 279 TAACATTAGCAGGTGTGATGACGAAAACTGTCTTACTGTCTCAATCATCGAAACA 338

QY 387 CAACACTGGTCTCAAGGCTGAAGAAATCATCAAGTTGATGAGGTGATTGAAGCTGCAAA 446
DB 339 TGATACTGGAGTTCAAGCTGAAAAA---CTCGACTTGGATGCTGTGTGAGCACTTGGCTAG 395
QY 447 GGAGGCAAGGCCAATGGATCTACAAGGTTCTGTATGGTCTGCTCTGGAGAGAGATGAA 506
DB 396 AGATGCGAAGCAGAGGTGTGACTAGAGTTTGGCTTGGCTTGGAGAGATATGCA 455
QY 507 AGGTAGAAAGTCAAACTTTGAAGAAATCAAAAGAGATGATCACTGCTGCTCCATGAC---CT 563
DB 456 AGGAAGAAGGCTCGCTTAGACAAAATCAGTAAAGTTGTTTAAACAAATCAACGACGAGTT 515
QY 564 TGGAAATGGAGAGTTGTGTCCCTTGGGAATGGTTGATTAAGACCAAGCCACTGAATTGAA 623
DB 516 GAAGTTGGAAACTTGTGTACTCTCGGGATGATCAACGAAAGAGCGGTGAAATTTTAA 575
QY 624 AAGTGTGGGTTGACGGGTACCAACCACTTAACATTAAGTACTTACAAAGGAACACTATCCAAA 683
DB 576 ACAGATGGTTGACTGCTTACCAACCACTTGAACACTTCAAGAGAACATTTATCCAAA 635
QY 684 GGTGATCTCCAAGAAAGCTTTTGATGATAGATTGAAACCACTTCAAAAACGTTTCAAGGATC 743
DB 636 CGTCGTCAACCAACCGGTATACAGCAAGAGATTAGAAACCACTTAAACACGTTCCAAAAGC 695
QY 744 TGGATTAAAGCATGACAGGTGTGATTTCTTGGTCTTGGTGAGACCCCAAGAGACCGTGT 803
DB 696 TGGTATCAAAAGCTTGTACCGGTGTTTCTCGGATTTGGGAGAAAACCGCACAGATCATGT 755
QY 804 ATCTTCTCTACACCTTGGCCCAATGGATCAGATCCAGAGTCTCTTCCAATCAACAG 863
DB 756 TTGTTTTTGTACACTTTGTCCCAATGAGTCCGCAACCACTTGGCAATCAACAG 815
QY 864 ACTGGTCCCAATCAAGGGCAGCGCAATGTATGAA-----GAAAGTTAAGAAACA 911
DB 816 ATTAGTTCCAATCCAGGGTACTCCAATGTTGAAAGTGTGAAAATGCACCAAGGAAAG 875
QY 912 GCAAGTTGAAGTTGATGAGATTGTGAGAACCACTTGTCTACTGCAAGATTGGTCTGCCAAC 971
DB 876 ACRAATGGAATTTGATGCTATTTTGAAGAACTATTGTCTACTGTAGATTGATTGCTCTGA 935
QY 972 GTCTATTATCAGATTGGCTCCAGGAAGATATACAAATCAAGAGGCGACAGAGGTGATGTG 1031
DB 936 ATCCATCATTAGATTGGCCCGCGGTCTCATACCATGAAGGAAACACGAGCAAGTCTTGTG 995
QY 1032 CTTTCATGGCTGGTTGTAATGCCATCTTCAAGGTTAAGAAAATGCTCAACAATGTGTAA 1091
DB 996 TTTCATGAGTGGGTCAATGCTATTTTACCGGTAAAGAGAAATGTTGACTCAATGTGTAA 1055
QY 1092 CGCTGGGATGAGATTAAGCCATGTTGGCTTAATGGGGTCTGAAACCAATGGAGAGTTT 1151
DB 1056 TGGTTGGGATGAAGATATCGCCATGTTTGAAGAAATGGGGTTTGAACCAATGGAAAGTTT 1115
QY 1152 CA 1153
DB 1116 CA 1117

RESULT 3
ABAO1978

ID ABA01978 standard; cDNA; 1659 BP.

XX ABA01978;

DT 06-FEB-2002 (first entry)

DE Soybean biotin synthase coding sequence #4.

XX Soybean; biotin synthase; biotin biosynthesis; herbicide;
KW plant development; transgenic plant; ss.

OS Glycine max.

XX Key

Location/Qualifiers

FT	CDS	47. .1300	
FT		/*tag= a	
FT		/product= "biotin synthase"	
XX			
XX	PN	US2001039042-A1.	
XX	PD	08-NOV-2001.	
XX	PF	19-DEC-2000; 2000US-00740288.	
XX	PR	21-DEC-1999; 99US-0172929P.	
XX	PA	(ALLE/) ALLEN S M.	
XX	PA	(KINN/) KINNEY A J.	
XX	PA	(MIAO/) MIAO G.	
XX	PA	(OROZ/) OROZCO E M.	
XX	PI	Allen SM, Kinney AJ, Miao G, Orozco EM;	
XX	PI	WPI: 2002-040723/05.	
XX	DR	P-PSDB; AAMS1985.	
XX	PT	New polypeptides, useful as targets for herbicide discovery, and as	
XX	PT	probes for genetic and physical mapping of genes of which they are part,	
XX	PT	or creating transgenic plants, comprises biotin synthase polypeptides and	
XX	PT	encoding polynucleotides.	
XX	PS	Claim 5; Page 35; 46pp; English.	
XX	CC	The present invention provides the protein and coding sequences of biotin	
XX	CC	synthase enzymes from barley, wheat, maize, soybean and the prickly	
XX	CC	poppy. These sequences can be used to produce transgenic plants which	
XX	CC	express different levels of the gene, or express it at different times in	
XX	CC	plant development. They can also be used as a target in the production of	
XX	CC	herbicides. The present sequence is a soybean biotin synthase cDNA	
XX	SQ	Sequence 1659 BP; 496 A; 334 C; 402 G; 427 T; 0 U; 0 Other;	
XX		Query Match 29.4%; Score 349.2; DB 6; Length 1659;	
XX		Best Local Similarity 60.6%; Pred. No. 1.6e-92;	
XX		Matches 595; Conservative 0; Mismatches 378; Indels 9; Gaps 1;	
QY	173	AAATGTCATGTCAGTGAAGACACAGTCAACACTCGACCCAAAGAGAAATTAAGCTA	232
DB	234	AAGAGCACTCAAGCACTAAGTCCCGCAGCAATGGACAGAGAGAGAAATCAAGGGA	293
QY	233	TATATGACACACCACTCATGAGCTTGTATGCTATGCTCAGGTGCAACAGAGAGTTCC	292
DB	294	TCTATGATTAAGCCATTGATGAGTATGTTGGGGTGTGTAGTTGACAGGAAATTC	353
QY	293	AACACCTTCAGAGTTCAATTTGCACTCTTATGAATATCAAACTGGTGTGTACCG	352
DB	354	ATATACCTGGGGCTATTTCAGATGTGTACATTTGTTGAACATCAAGACGGGTGTTCTCGG	413
QY	353	AGGACTGTAGTACTGTGCCCCAATCAACAGCTTCAACACTGGTGTCAAGGCTGAAAGAA	412
DB	414	AGGATTGTTCTTACTGGCCCAATCATCCGCTACCAACCGGTCTCAAGCTCCAAAA	473
QY	413	TCAATCAAGTTGATGAGGTGATTAAGCTGCAAAAGGAGGCAAGCCAAATGGATCTACAA	472
DB	474	TGGTCTCCGTGCAATCTGTCTCTCAGCGCCGCGCATCGCCAAAGACAACGGTAGTACAC	533
QY	473	GGTTCGTATGGGTGCTGTTGGAGAGAGATGAAGGTAGAACTCAAACTGAAGAAA	532
DB	534	GTTCCTGATGGGAGCGCGTGGCGCATATGCTGGACGAAACCAATCTCAAATG	593
QY	533	TCAAAGAGATGATCACTGCTGTCTCCATGACCTTGAATGGAGAGTTGTCAACCTGGGAA	592
DB	594	TCAAACAATGGTTAGCGAGATTCGCGGAATGGGTATGGAGTATGTGTACGCTTGTA	653
QY	593	TGGTTGATAAGACAAGCCCACTGAATTTGAAAAGTGTGGTTGACGGCGTACACATA	652
DB	654	TGATTGATGACAGCAAGCTCAGGAACTCAAAAGAGCGGTCTCACGGCTTATAATCAT	713
QY	653	ACATTGATCTTACAAGGAACACTATCCAAAGGTGATCTCCAAAGAAGCTTTGATGATA	712
DB	714	ATGTGATAGTTCGAGGGATTTCTATCCCAAGGTATACGACACAGGACTTATGATGAGA	773
QY	713	GATTGAAAACATTCAAAAGCTTCAAGGATCTCGATTAAGAGGATGACAGAGGTGATTC	772
DB	774	GATTGGATACCAATTAAGAAATGTGAGAGAGCGCGGAATCAATGTTGTACGGGTGAATCC	833
QY	773	TTGGTCTTGGTGAGACCCCAAGAGACCGTGATATCTTTCTCTACACCTTCGCCCAATGG	832
DB	834	TCGATTAGGAGAAATAAGTCTGACCATATTTGGACTTTTGGAGACGGTTGCTACGTTC	893
QY	833	ATCAGCATCCAGAGTCTCTTCCCAATCAACAGACTGGTCCCAATCAAGGGCAACGCAATGT	892
DB	894	CTTCGCATCGGAATCATTTCTCTGTGAACATGTTAGTGGCTATCAAGGAACACCACTGG	953
QY	893	ATGAAGAAGTTAAGAACAGCAAGTTGAGTTGATGAGATGTGCAGAACCATTTGCTACTG	952
DB	954	AAGGA-----AACAGAGAGGTGGAATTTGAGAAATATGTTGAGAAATGTTTGCACGG	1004
QY	953	CAAGATTGGTTCATGCCCAACGCTATTATCAAGATTGGCTGCAGGAAGATATACAAATGAAG	1012
DB	1005	CTAGAATCGTCATGCCCTAAACCATCGTGGCTTTGGCAGCTGGAAGAGGAGAAATTGACG	1064
QY	1013	AGCAGAACACAGGTGATGCTTCATGGCTGGTTGTAATGCCATCTTTCACAGGTAAGAAA	1072
DB	1065	AGGAACAACAGGTCTTATGTTTCATGGCGGAGCAATGCGGTTTTCACAGGAGAAACAA	1124
QY	1073	TGCTCAACAATGTGTAAACGGCTGGATGAGGATAAGCCATGTTGGCTAAATGGGGTC	1132
DB	1125	TGTTAAACCAACACAGCGCTTGGATGGGTGTCGATTCGCTCGGTTTTCACAGATGGGGAT	1184
QY	1133	TGAACCAATGGAGAGTTTCAA 1154	
DB	1185	TAAGACCCATGGAAAGTTTCGA 1206	
XX	RESULT 4		
XX	ADE76391		
ID	ADE76391	standard; cDNA; 1659 BP.	
AC	ADE76391;		
XX	AC		
DT	29-JAN-2004	(first entry)	
XX	DE	Soybean SID30 biotin synthase cDNA - SEQ ID 29.	
XX	XX	biotin synthase; plant; breeding; seed; herbicide; ss; gene; soybean;	
XX	KW	SID30.	
XX	OS	Glycine max.	
XX	PN	US2003192073-A1.	
XX	PD	09-OCT-2003.	
XX	PF	07-MAY-2003; 2003US-00431544.	
XX	XX		
PR	21-DEC-1999;	99US-0172929P.	
PR	19-DEC-2000;	2000US-00740288.	
XX	XX	(ALLE/) ALLEN S M.	
PA	(KINN/) KINNEY A J.		
PA	(MIAO/) MIAO G.		
PA	(OROZ/) OROZCO E M.		
XX	Allen SM, Kinney AJ, Miao G, Orozco EM;		
XX	WPI: 2003-831831/77.		
DR	P-PSDB; ADE76392.		
XX	PT	New isolated polynucleotide encoding plant biotin synthase, useful in	

erythropoietin, tissue-type plasminogen activator, synthetic haemoglobin and insulin. They can be used in applications, such as separation matrices (e.g. dialysis membranes). Collagen may be combined with the peptides to produce membranes for use as artificial skin. The MM may be used for making very thin, transparent fabric. Drugs which inhibit the self assembly of the peptides into filaments or filamentous membranes may be useful for treating Alzheimer's disease or scrapie infection. As they are resistant to proteolytic digestion and alkaline and acidic pH (such as stomach acid), drug delivery vehicles made of the MM could be taken orally. The charged residues and conformation of the MM promote cell adhesion and migration. The permeability of the MM also permits diffusion of small molecules, to the underside of cell monolayers, useful for tissue culture of differentiated cells and/or stratified cell layers. (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

Query Match 27.6%; Score 328.2; DB 2; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

QY 553 GTCATGACCTTGAATGGAGAGTTGTGTACCCCTGGGAATGTTGATTAAGCCAGCC 612
DB 6 GTGAATGATATGGGGCTAGAAACGTTGTACTTGTAGTATGTTGATCAAGATCAAGCA 65
QY 613 ACTGAATTGAAAGTCTGGTGTGACGGCTGACACCATTAACATTTGATTAACAGGAA 672
DB 66 AAGCAATTAAGATGAGGTTGTGACTGCTGATACACCATTAACATTCGACATTCAGAGAA 125
QY 673 CACTATCCAAAGGTGATCTCCACAGAAAGCTTTGATGATAGATTGAAACATTCAGAAAC 732
DB 126 CACTATAGTAGTTCATCACACGAGAACTTACGACAGAGTTTACAGACCATCAAGAAAT 185
QY 733 GTTCAAGGATCTGAATTAAGGATGACACAGGTGTTCTTGGTCTTGGTGAGACCAA 792
DB 186 GTCCAGAAATCTGAATTAAGGCTGTACCGGTGGTATTTGGGTCTCGGTGAAGCGAA 245
QY 793 GAGGACCGTGTATCTTCTCTACACCTTGGCCACATGATCAGATTCAGAGTCTCTT 852
DB 246 GAGCACCATATAGATTATCTATACATATTCATATGCTCTCTATCTCTGATGCCCTA 305
QY 853 CCAATCAACAGACTGTTCCCAATCAAGGGCACGCCAATGTATGAAGAAATT----- 903
DB 306 CCAATTAATAGACTAGTTGCTATCAAGGGACTTCCATGCTGAGGAACTTGGCGATCCA 365
QY 904 AAGACAGCAAGTTGAAGTTGATGATGATGTCAGAACCAATTCGACTGCAAGATTGGTC 963
DB 366 AAGAGTAAAGAAAGTTGCAATTCGACGAAATTTGAGAACCAATTCGACGCGAGATAGTT 425
QY 964 ATGCCAACGCTTATATCAGATTGGCTGACGAGAGATATACATGAAGAGGCGAGAACAG 1023
DB 426 ATGCCAAGGCCATTATAGACTTGGCTGCTGCTGTTATACATGAAGAAACAGAGCAA 485
QY 1024 GTGATGTGCTTCATGGCTGGTGTGTAATGCCATCTTACAGGTGAAGAAATGCTCACAACA 1083
DB 486 TTTGTCTGTTTCATGGCAGGTTGTAACAGTATCTTACCCGTAAGAAATGCTGACGACA 545
QY 1084 ATGTGTAACGGCTGGGATGAGGATTAAGCCATGTTGCTTAATGGGTCTGAACCAATG 1143
DB 546 ATATATAACGGTTGGGACGAGACAGGCAATGTTGGCTTAATGGGGATGCAACCTATG 605
QY 1144 GAGAGTTTCAAAATACAA 1160
DB 606 GAGGCATTAAGTACGA 622

RESULT 6
AAZ11218
ID AAZ11218 standard; DNA; 3083 BP.
XX
AC AAZ11218;
XX
DT 09-NOV-1999 (first entry)

XX Human zuotin coding sequence.
DE Membrane forming peptide; cell culturing; macroscopic membrane;
XX amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
KW artificial skin; separation matrix; artificial tissue; scrapie infection;
KW Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
KW protein conformational disease; human; zuotin; ZUO1; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1292..2593
FT CDS /tag= a
FT /product= "zuotin"
XX
XX US955343-A.
XX 21-SEP-1999.
XX 22-AUG-1994; 94US-00293284.
XX 28-DEC-1992; 92US-00973326.
XX (NASI) MASSACHUSETTS INST TECHNOLOGY.
XX Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;
XX WPI; 1999-539576/45.
XX P-PSDB; AAY32954.
XX Cell cultures utilizing stable macroscopic membranes formed by the self-
XX assembly of amphiphilic peptides.
XX Disclosure; Col 33-38; 49pp; English.

XX This sequence encodes the human zuotin (ZUO1) protein. Fragments of the
XX protein can be used in the in vitro method of the invention. The method
XX is for culturing cells utilising stable macroscopic membranes formed by
XX the self-assembly of amphiphilic peptides. The peptides are salt-induced
XX to form insoluble and protease-resistant protein filaments with a beta-
XX sheet secondary structure. The membranes may be useful in a wide
XX variety of medical, research, industrial and biomaterial applications
XX such as slow-diffusion drug delivery systems, artificial skin and
XX separation matrices. The membranes may be used to support in vitro cell
XX attachment and growth and for supporting artificial tissue (e.g. for in
XX vivo use as implants). They are particularly useful as experimental
XX models for Alzheimer's disease and scrapie infection and so may be used
XX in disease modelling experiments and to assay for agents which modulate
XX the disease processes. Additionally, they may be used in this way to
XX study liver cirrhosis, kidney amyloidosis and other protein
XX conformational diseases. The membrane in the membrane/cell mixture
XX produced in the method: (i) supports cellular attachment and growth; (ii)
XX self-assembles to form large, macroscopic membranes that are insoluble
XX and stable in aqueous solutions, serum and ethanol; (iii) is highly
XX resistant to heat, alkaline/acidic pHs, chemical denaturants and
XX proteolytic digestion; (iv) is non-cytotoxic and non-immunogenic; (v) is
XX visible to the naked eye if dyed but is otherwise transparent; (vi) may
XX form thin, permeable, high density sheets or fibril like structures with
XX simple structures, high tensile strength and a porous structure; (vii)
XX may be metabolised by humans and animals; (viii) is inexpensive to
XX produce; and (ix) can be produced and stored in sterile conditions
XX
XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

Query Match 27.6%; Score 328.2; DB 2; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
QY 553 GTCCATGACCTTGAATGGAGAGTTGTGTACCCCTGGGAATGTTGATTAAGCCAGCC 612
DB 6 GTGAATGATATGGGGCTAGAAACGTTGTACTTGTAGTATGTTGATCAAGATCAAGCA 65


```
Db 246 GAGGACCATATAGGATTCATCTACACATTATCCAAATATGTCTCTCATCTCGATCCCTA 305
Qy 853 CCAATCAACAGACTGTGCCAATCAAGGCGACGCCAATGATGAAGAAGTT----- 903
Db 306 CCAATTAAATAGACTAGTGTCTATCAAGGGAGCTCCAATGGCTGAGGAACCTTGGCGATCCA 365
Qy 904 AAGAACCAAGCAAGTTGAAGTTGATGAGATTGTGAGAACCAATGCTACTGCAAGATTGGTC 963
Db 366 AAGAGTAAAAAGTTGCAATTCGACGAAATTTTGAGAACCAATTCACACAGCGAGAATAGTT 425
Qy 964 ATGCCAACGCTTATTATCAGATTGGCTGCGAGGAAGATATACAAATGAAGAGGCGAGAACAG 1023
Db 426 ATGCCAAAGGCCAATTATAGACTTGCCTGCTGGCTGCTGCTTATACAAATGAAGAAACAGAGCAA 485
Qy 1024 GTGATGTCTTTCATGGCTGTGTTGAATGCCATCTTACAGCTAAGAAAATGCTCAACA 1083
Db 486 TTTGCTGTTTTCATGGCAGGTTGTAACAGTATCTTACCCGTAAGAAAATGCTGACGACA 545
Qy 1084 ATGTGTAACGGCTGGGATGAGGATAAAGCCATGTTGGCTAAATGGGGTCTGAAACCAATG 1143
Db 546 ATATATAACGGTTGGGACGAGAACAGGCAATGTTGGCTAAATGGGGATTGCAACCTATG 605
Qy 1144 GAGAGTTTCAAAATACAA 1160
Db 606 GAGGCATTTAAGTACGA 622
```

RESULT 8

```
AD004472
XX ID AD004472 standard; DNA; 3083 BP.
AC AD004472;
DT 29-JUL-2004 (first entry)
XX Yeast Z-DNA binding protein (zuotin) DNA.
DE Macroscopic membrane; cell growth; EAK16; Z-DNA binding protein; zuotin;
KW medical product; suture; artificial skin; internal lining;
KW slow-diffusion drug delivery system; protein-type drug; erythropoietin;
KW haemoglobin; insulin; cell adhesion; cell migration; Alzheimer's disease;
KW scrapie infection; yeast; gene; ds.
XX Saccharomyces cerevisiae.
PH Key Location/Qualifiers
FT repeat_region 701..714
FT /*tag= b
FT /rpt_type= TANDEM
FT repeat_unit 701..702
FT /*tag= a
FT /note= "7 AT repeats"
FT repeat_region 850..859
FT /*tag= c
FT /rpt_type= TANDEM
FT repeat_region 950..958
FT /*tag= d
FT /rpt_type= TANDEM
FT repeat_region 1016..1031
FT /*tag= f
FT /rpt_type= TANDEM
FT repeat_unit 1016..1017
FT /*tag= e
FT /note= "8 AT repeats"
FT repeat_region 1153..1162
FT /*tag= g
FT /rpt_type= TANDEM
FT repeat_region 1292..1293
FT /*tag= h
FT /products= "Yeast Z-DNA binding protein (zuotin)"
FT misc_feature 2159..2206
FT /*tag= i
FT /note= "Non-purine tract"
```

```
XX US2004087013-A1.
PN 06-MAY-2004.
XX 17-MAR-2003; 2003US-00390472.
XX 28-DEC-1992; 92US-00973326.
PR 22-AUG-1994; 94US-00293284.
PR 26-MAR-1997; 97US-00824515.
XX (HOLM/) HOLMES T.
PA (ZHAN/) ZHANG S.
PA (RICH/) RICH A.
PA (DIPE/) DIPERSIO C M.
PA (LOCK/) LOCKSHIN C.
XX Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
PI WPI; 2004-356208/33.
XX P-PSDB; ADO04473.
XX Novel EAK16 protein incorporated into macroscopic membranes, useful in
PT biomaterial applications such as medical products, artificial skin or
PT internal linings, slow-diffusion drug delivery systems for in vitro cell
PT growth.
XX Claim 30; SEQ ID NO 1; 56pp; English.
XX The invention relates to a method for in vitro cell culture which
CC involves adding a macroscopic membrane that is formed by self-assembly of
CC amphiphilic peptide in an aqueous solution containing monovalent metal
CC cations to a cell culture medium comprising cells, thus forming a
CC membrane/culture mixture, maintaining the mixture under conditions
CC sufficient for cell growth. The invention also relates to EAK16 peptide
CC derived from yeast Z-DNA binding protein (zuotin). Zuotin incorporated
CC into the macroscopic membranes are useful in biomaterial applications
CC such as medical products (e.g., sutures), artificial skin or internal
CC linings, slow-diffusion drug delivery systems supports for in vitro cell
CC growth or culture and support for artificial tissue for in vivo use, as
CC slow-diffusion drug delivery vehicle for delivering protein-type drugs
CC e.g., erythropoietin, synthetic haemoglobin, insulin, etc., useful as
CC conductive biopolymer for culturing cell monolayers, for promoting cell
CC adhesion and migration, useful as experimental models for Alzheimer's
CC disease and scrapie infection. The present sequence is yeast zuotin DNA.
CC This sequence is used to illustrate the method of the invention.
XX SQ Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;
```

```
Query Match 27.6%; Score 328.2; DB 12; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
Qy 553 GTCATGACCTTGGAAATGGAGAGTGTGTGCACCTCGGGAAATGGTTGATAAGACCAAGCC 612
Db 6 GTGAATGATATGGGCTAGAAAACGTTGTTACTTTAGGTATGGTTGATCAAGATCAAGCA 65
Qy 613 ACTGAATTCAAAAGTCTGGGTTGACGGGTACACCAACATAACATTGATCTACTACAGGAA 672
Db 66 AAGCAATTGAAAAGATGCGAGGTTTGACTGTCATACAAACCATACATCGCACTTCCAGAGAA 125
Qy 673 CACTATCCAAAGGTGATCTTCCAAAGAGCTTTGATGATAGATTGAAAAACATTCAAAAAC 732
Db 126 CACTATAGTAAGTTCATCACACAGAGAACCTACGACGACAGGTTACAGACCATCAAGAT 185
Qy 733 GTTCAAGGATCTGGATTAAGGCATGCACAGGTGGTATTTCTTGGTCTTGGTGAGACCCAA 792
Db 186 GTCCAAGAATCTGGAATRAAAGCCTGTACCGGTGGTATTTTGGGTCTCGGTGAAAGCGAA 245
Qy 793 GAGGACCGGTATCTTTCTCTACACCTTGGCCACATGGATGCAGATCCAGATCTCTT 852
Db 246 GACGACCATATAGGATTCATCTACACATTATCAATATGTTCTCTCATCTGATGCCCTA 305
```



```

Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGCGACGCCCAATGTATGAAGAAGTT----- 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 306 CCAATTAATAGACTAGTTGCTATCAAGAGGACTCCCAATGGGTGAGGAACCTTGGCGATCCA 365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 904 AAGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCAATTGCTACTGCAAGATTGCTC 963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 366 AAGAGTAAAGTTGCAATTGACGAAATTTGAGAACCAATTGCGACAGCGAGATAGTT 425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 964 ATGCCAACGCTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAGAGCGCAGAACAG 1023
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 426 ATGCCAAAGGCCAATATAGACTTGGCGTGGTCTGTATACATGAAAGAACAGAGCAA 485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1024 GTGATGCTTCATGCGTGGTGTGTAATGCCATCTTTCACAGGTAAGAAATGCTCACACA 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 486 TTTGCTCTGTTTCATGGCAGGTTGTAAACAGATATCTTCACCGGTAAAGAAATGCTACGACA 545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1084 ATGTGTACGCTGGATGAGGATGAAGCAAGCCATGTTGGCTAAATGGGTCTGAACCAATG 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 546 ATATATAACGGTGGGACGAGACAAAGGCAATGTTGGCTAAATGGGGATTGCAACCTATG 605
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1144 GAGAGTTTCAAAATACAA 1160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 GAGGCATTTAAGTACGA 622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
ADSI5330
ID ADS15330 standard; DNA; 3083 BP.
AC ADS15330;
XX
XX 16-DEC-2004 (first entry)
XX
XX Yeast DNA binding protein zootin DNA.
XX
XX cell culture; macroscopic membrane; amphiphilic peptide; biomaterial;
XX slow-diffusion drug delivery system; artificial skin; separation matrix;
XX Alzheimer's disease; scrapie; liver cirrhosis; kidney amyloidosis;
XX protein conformational disease; yeast; DNA binding protein; zootin; ds;
XX gene.
XX
XX Ascomycota.
XX
XX Key Location/Qualifiers
XX CDS 1292..2593
XX FT /*tag= a
XX FT /product= "Zootin"
XX
XX US6800481-B1.
XX
XX 05-OCT-2004.
XX
XX 26-MAR-1997; 97US-00824513.
XX
XX 28-DEC-1992; 92US-00973326.
XX
XX 22-AUG-1994; 94US-00293284.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
XX WPI; 2004-707224/69.
XX
XX P-PSDB; ADS15331.
XX
XX Culturing cells, in vitro, by adding macroscopic membrane formed by self-
XX assembly of amphiphilic peptides and monovalent metal cations, to form
XX culture mixture, and maintaining culture mixture under conditions for
XX cell growth.
XX
XX Example 5; SEQ ID NO 1; 50pp; English.
XX
XX The invention describes a method of culturing (M1) cells, in vitro. The
XX cell culture involves adding a macroscopic membrane which is formed by

```

```

CC self-assembly of amphiphilic peptides in an aqueous solution containing
CC monovalent metal cations, where the peptides have alternating hydrophobic
CC and hydrophilic amino acids and are complementary and structurally
CC compatible, to a cell culture medium comprising cells, thus forming a
CC membrane/culture mixture, and maintaining the mixture under conditions
CC sufficient for cell growth. (M1) is useful for culturing a macroscopic
CC membrane utilised in biomaterial applications e.g., slow-diffusion drug
CC delivery systems, artificial skin or separation matrices, or as
CC experimental models for Alzheimer's disease and scrapie infectional
CC liver cirrhosis, kidney amyloidosis, or other protein conformational
CC diseases. (M1) enables in vitro culturing of macroscopic membrane that is
CC stable in aqueous solution, serum and ethanol, highly resistant to heat,
CC alkaline and acidic pH, chemical denaturation and proteolytic digestion,
CC and is non-cytotoxic. This sequence encodes yeast DNA binding protein
CC zootin from which amphiphilic peptides capable of forming membrane can be
CC isolated.
XX
XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

```

```

Query Match 27.6%; Score 328.2; DB 13; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

```

```

Qy 553 GTCCATGACCTTGAATGGAGATTGTGTACCCCTGGGAATGGTTGATAAGACCAAGCC 612
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 613 ACTGAATTGAAAAGTGTGGTTCACGGCGTACAAACCAATACATTGTACTTTACAAGGAA 672
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 66 AAGCAATTGAAGATGAGGTTTGCTGCTACAAACCAATACATCGACACTTCGAGANA 125
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 673 CACTATCCAAAGGTGATCTCCCAAGAAAGCTTTTGATGATAGATTGAAACAATTCAAAAC 732
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 126 CACTATAGTAGTCAATCACCACGAGAACCTTACGACGACAGTTTACAGCCATCAAGAT 185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 733 GTTCAAGATCTGAATTAAGGCATGACAGGTGGTATTCTTGGTCTTGGTGAGCCCAA 792
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 186 GTCCAAGAAATCTGAATAAAGCCCTGTACCGGTGGTATTTTGGGTCTCGGTGAAGCGAA 245
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 793 GAGGACCGTGTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTT 852
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 246 GACGACCATATAGGATTCATCTACACATTTATCCAAATATGTCTCTCATCTCGAGTCCCTA 305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGGCAACGCAATGTATGAAGAAGTT----- 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 306 CCAATTAATAGACTAGTTGCTATCAAGGGACTCCCAATGGCTGAGGAACCTTGGCGATCCA 365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 904 AAGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCAATTGCTACTGCAAGATTGCTC 963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 366 AAGAGTAAAGTTGCAATTGCGACGAAATTTTGAAGAACCAATTTGCCACAGCGAGATAGTT 425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 964 ATGCCAACGCTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAGAGCGCAGAACAG 1023
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 426 ATGCCAAAGGCCAATTAAGACTTGGCGTGGTCTGTATACATGAAAGAACAGAGANA 485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1024 GTGATGCTTCATGCGTGGTGTGTAATGCCATCTTTCACAGGTAAGAAATGCTCACACA 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 486 TTTGCTCTGTTTCATGGCAGGTTGTAAACAGATATCTTCACCGGTAAAGAAATGCTACGACA 545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1084 ATGTGTAAACGGCTGGGATGAGGATGAAGCCATGTTGGCTAAATGGGTCTGAACCAATG 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 546 ATATATAACGGTGGGACGAGACAAAGGCAATGTTGGCTAAATGGGGATTGCAACCTATG 605
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1144 GAGAGTTTCAAAATACAA 1160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 GAGGCATTTAAGTACGA 622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 10
ABAO1974
ID ABA01974 standard; cDNA; 1515 BP.
XX
XX ABA01974;
AC

```


XX 06-FEB-2002 (first entry)
 DT Maize biotin synthase coding sequence #5.
 DE
 XX
 XX Maize; biotin synthase; biotin biosynthesis; herbicide;
 KW plant development; clone ch0c.pk009.j14; transgenic plant; ss.
 KW
 XX Zea mays.
 XX
 XX Key Location/Qualifiers
 FT 131..1264
 CDS /*tag= a
 FT /product= "biotin synthase"
 FT
 PN US2001039042-A1.
 XX
 XX 08-NOV-2001.
 PD
 XX 19-DEC-2000; 2000US-00740288.
 XX
 XX 21-DEC-1999; 99US-0172929P.
 PR
 XX (ALLE/) ALLEN S M.
 PA (KINN/) KINNEY A J.
 PA (MIAO/) MIAO G.
 PA (OROZ/) OROZCO E M.
 XX
 XX Allen SM, Kinney AJ, Miao G, Orozco EM;
 PI
 XX WPI: 2002-040723/05.
 DR
 DR P-PSDB; RAMS1981.
 XX
 XX New polypeptides, useful as targets for herbicide discovery, and as
 PT probes for genetic and physical mapping of genes of which they are part,
 PT or creating transgenic plants, comprises biotin synthase polypeptides and
 PT encoding polynucleotides.
 XX
 XX Claim 5; Fig 2; 46pp; English.
 PS
 XX The present invention provides the protein and coding sequences of biotin
 CC synthase enzymes from barley, wheat, maize, soybean and the prickly
 CC poppy. These sequences can be used to produce transgenic plants which
 CC express different levels of the gene, or express it at different times in
 CC plant development. They can also be used as a target in the production of
 CC herbicides. The present sequence is a maize biotin synthase cDNA obtained
 CC from clone ch0c.pk009.j14, which was derived from a maize embryo
 XX
 XX Sequence 1515 BP; 467 A; 368 C; 368 G; 312 T; 0 U; 0 Other;
 SQ
 Query Match 26.2%; Score 311.8; DB 6; Length 1515;
 Best Local Similarity 59.1%; Fred. No. 2e-81;
 Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;
 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
 251 AACACTGGAGCGGCCGAGATCCAGCGCTCTACGACTCACCGCTCTCGACTCCTC 310
 262 CACTATGCTCAGGTGCACACAGAGGTTCCAAACACCTTCAGAGGTTCAATGTGCACT 321
 311 TTTCAGGGGCTCAGGTCCACAGAAATGTCATAAATTCAGAGAAATGCAGCAATGCACA 370
 322 CTTATGATATCAAACTGGTGTGTACCGGACTGTAAAGTACTGTGCGCAATCACAG 381
 371 CTTCTTTTCAATCAAGACTGGTGGATGTCAGTGAAGATTGTTCTTACTGTCTCAGTCATCA 430
 382 CGTTACACACTGCTGTCAAGCTGAAAGAAATCATCCAAAGTTGATGAGGTGATTGAAGCT 441
 431 AGATACACACTGATTAAGGCCCAAAATTTGATGACAAAGATGCTGTCTTGGAGCA 490
 442 GCAAAGGAGGCAAGGCCAATGGATCTTACAAGGTTCTGTATGGGTGCTGTGGAGAGAG 501
 491 GCAAAAAGGCAAAAGAGTCTCTGGGAGCACCGGTTTTTTCATGGGAGCTGTCATGGAGAA 550

502 ATGAAGGTAGAAAAGTCAAACTTTGAAGAAAATCAAGAGATGATCACTGCTGTCTCCATGAC 561
 551 ACCATTGGCAGGAAATCAAACTTTCAACAGATTTCTTGAATATGTCAAGGAAATTAAGGGT 610
 562 CTTGGAATGAGAGATTGTGTCACCTGGGAATGGTGTATTAAGACCAAGCACCCTGAATTG 621
 611 ATGGCATGGAGGTCTGTGTCACACTAGGATGATAGAGAAACAACAAGCTGAAGAACTC 670
 622 AAAAGTCTGGGTTGACGGGTACAACATAAATTGATATCTTACAAGGAAACATATCCA 681
 671 AAGAGGCTGGACTTACAGCATATATCATACCTAGATATCATCAAGAGATATATCCC 730
 682 AAGGTGATCTCCACAAGAGCTTTTGATGATAGATTGAAAAATTCAAAAAGCTTCAAGGA 741
 731 AACATTATTACCACAAGATCATATGATGATAGACTGTCAGACTCTTTGAGCATGTCCGTGAA 790
 742 TCTGGATTAAGGATGACACAGGTGGTATTTCTGGTCTTGGTGAGACCAAGAGACCGT 801
 791 GCTGGAATAAGCATCTGCTCAGGTGGAATCAITGGTCTTGGTGAAGCAGAGGAGCCGG 850
 802 GTATCTTTCTCTACACCTTTGGCCACAATGATGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
 851 GTAGGTTGTTGCATACCTTAGCTACCTTCCCTACACCCAGAGAGCGTTCTTATTAT 910
 862 AGACTGCTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTTAAAGAAACAAGCAAGTTGAA 921
 911 GCATTGGTGTCTGTAAGGACACACTCTTTGAGG-----ACCAGAGCCTGTAGAG 961
 922 GTTGATGAGATTGTGAGAACCATTTGTAATCTGCAAGATTGTTGATGATGATGATGATGAT 981
 962 ATCTGGGAAATGATCCGATGATCGCCACTGCTCGGATCAGATGATGATGATGATGATGAT 1021
 982 AGATTGGCTCAGGAGATATACATGAAGAGGAGCAGAGGATGATGATGATGATGATGATGAT 1041
 1022 AGGCTTTTCAAGAGGCGGAGTACCGTTCTCGATGCAAGAAAGAGGCTGTGCTTCTCGCT 1081
 1042 GGTGTAATGCCATCTTCAAGGTAAGAAAATGCTCAACAATGTTGTAAGCGGCTGGAT 1101
 1082 GGGGCCAATCCATCTTTGCGGCGGAGAACTTCTCAACCCGCAAAACAGACTTTTGTAT 1141
 1102 GAGATAAAGCCATGTTGGCTTAATGGGGTCTGAACCAATGG 1144
 1142 GCGAGCAGGCGATGTTCAAGATCCTTGGGCTGTATCCCAAGG 1184

RESULT 11
 ADE76383
 ID ADE76383 standard; cDNA; 1515 BP.
 XX AC ADE76383;
 XX AC ADE76383;
 XX 29-JAN-2004 (first entry)
 DT
 XX Corn SID22 biotin synthase cDNA - SEQ ID 21.
 DE
 XX biotin synthase; plant; breeding; seed; herbicide; ss; gene; maize; corn;
 KW SID22.
 KW
 XX Zea mays.
 OS
 XX
 XX US2003192073-A1.
 PN
 XX 09-OCT-2003.
 PD
 XX 07-MAY-2003; 2003US-00431544.
 PF
 XX 21-DEC-1999; 99US-0172929P.
 PR
 XX 19-DEC-2000; 2000US-00740288.
 XX
 XX (ALLE/) ALLEN S M.
 PA (KINN/) KINNEY A J.
 PA (MIAO/) MIAO G.

PA	(OROZ/) OROZCO E M.	
XX		
PI	Allen SM, Kinney AJ, Miao G, Orozco EM;	
XX		
DR	WPI: 2003-831831/77.	
DR	P-PSDB; ADE76384.	
XX		
PT	New isolated polynucleotide encoding plant biotin synthase, useful in	
PT	plant molecular biology and plant breeding purposes, particularly in the	
PT	production of altered levels of biotin synthase in plants and seeds.	
XX		
XX	Claim 5; SEQ ID NO 21; 47pp; English.	
XX		
CC	The invention relates to a novel isolated polynucleotide encoding a first	
CC	polypeptide or a third polypeptide where the first or third polypeptide	
CC	is a biotin synthase. The methods and compositions of the invention may	
CC	be useful in plant molecular biology and breeding, particularly in the	
CC	production of altered levels of biotin synthase in plants and seeds. The	
CC	polypeptides may be used as a target to facilitate the design and/or	
CC	identification of inhibitors of these enzymes that may be useful as	
CC	herbicides. Furthermore, the polynucleotides may be used as probes for	
CC	genetically and physically mapping the genes of which they are a part.	
CC	The current sequence is that of the corn biotin synthase-related CDNA of	
CC	the invention.	
XX		
SQ	Sequence 1515 BP; 467 A; 368 C; 368 G; 312 T; 0 U; 0 Other;	
	Query Match 26.2%; Score 311.8; DB 10; Length 1515;	
	Best Local Similarity 59.1%; Pred. No. 2e-81;	
	Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;	
QY	202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261	
Db		
QY	251 AACGACTGGAGCGCGCGAGATCCAGCGCTTACGACTCACGCTCTCGACTCTC 310	
Db		
QY	262 CACTATGCTCAGGTGCAACACAGAGGTTCCAAACCTTCAGAGTTTCAATTGTGCACT 321	
Db		
QY	311 TTTCCGCGGCTCAGGTCCACAGAAATGTCATAAATTCAGAGAAATGTCAGCAATGCACA 370	
Db		
QY	322 CTTATGAATATCAAACTGGTGGTTGACGAGGACTGTAAGTACTGTGCGCAATCACAG 381	
Db		
QY	371 CTTCTTTCAATCAAGATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 430	
Db		
QY	382 CGTTACAACACTGGTGTCAAGGCTGAAAGAAATCATCAAGTTGATGAGGTGATTTGAAGCT 441	
Db		
QY	431 AGATACAACACTGGATTGAAGGCCCAAAATGATGAACAAGATGCTGCTTGGAGCA 490	
Db		
QY	442 GCAAGAGGCAAGGCAATGATCTACAAGTTTCTGTATGGTGTCTGTGGAGAGAG 501	
Db		
QY	491 GCAAAAAGGCAAAAGAGTCTGGGAGCACCCGTTTTTGTGATGGAGTGTGATGGAGAGAA 550	
Db		
QY	502 ATCAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAAATCAAGAGATGATCTGTGTCATGAC 561	
Db		
QY	551 ACCATTGGCAGGAAATCAAACTTCAACCAAGTTCTTGAATATGTCAAGGAAATAAGGGT 610	
Db		
QY	562 CTTGGAATGAGAGTTGTGTCACTCGGGAATGGTGTATAAAGCAAGCACTGAATTG 621	
Db		
QY	611 ATGGCATGAGGCTGTGTCACACTAGGATGATGAGAGAAACAAGCTGAAGACTC 670	
Db		
QY	622 AAGGTGCTGGTTGAGCGGTAACAACATAAATTGATGATGATGATGATGATGATGATGATGATGAT 691	
Db		
QY	671 AAGAAGCTGGAATTACAGCATATATCAATAAAGCTAGATGATGATGATGATGATGATGATGATGATGAT 730	
Db		
QY	682 AAGGTGATCTCCACAGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 741	
Db		
QY	731 AACATTATTACCAAGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790	
Db		
QY	742 TCTGGATTAAAGGCATGCAGGTGGTATTCTTGGTCTTGGTGTGAGACCCCAAGAGGACCGT 801	
Db		
QY	791 GCTGGATTAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGTGAGAGGAGGACCGG 850	
Db		
QY	802 GTATCTTTCTTACACCTTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861	
Db		

Db	851 GTAGGGTTGTGCATACCCCTAGCTACCTTGCCTACACACCCAGAGAGCGTTCTTATTAAAT 910	
QY	862 AGACTGGTCCCAATCAAGGCGCAGCCCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921	
Db	911 GCATTGGTTGCTGTAAAGGCAACACCTCTTGAGG-----ACCAGAGCCTGTAGAG 961	
QY	922 GTTGATGAGATTGTGAGAACCAATTGTCTATGCAAGATTGGTCAATGCCAACGCTCTATTATC 981	
Db	962 ATCTGGGAATGATCCCGATGATCGCCACTGCTCGGATCAGATGCCAAAGCAATGGTG 1021	
QY	982 AGATTGGCTCAGAGAGATATACATGAAGAGGACAGACAGGATGATGCTTCTCATGGCT 1041	
Db	1022 AGGCTTTTCAGAGCGCGGAGTACGGTTCTCGATGCCAGACAGCGCTGTGCTTCTCGCT 1081	
QY	1042 GGTGTAATGCCATCTTCACAGGTAAAGAAATGCTCAACAATGTGTAAACGGCTGGGAT 1101	
Db	1082 GGGGCCAATCTCATCTTTGGCGGGAGAACTTCTCAACCCGCAACACAGACTTTGAT 1141	
QY	1102 GAGGATAAAGCCATTTGGGCTAAATGGGGTCTGAAACCAATGG 1144	
Db	1142 GCGGACCAAGCGATGTTCAAGATCCTTGGGCTGATGCCCAAGG 1184	
RESULT 12		
AAC48072		
ID	AAC48072 standard; DNA; 1313 BP.	
XX		
AC	AAC48072;	
XX		
DT	18-OCT-2000 (first entry)	
XX		
DE	Zea mays DNA fragment SEQ ID NO: 56164.	
XX		
KW	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic; pathway;	
KW	promoter; termination sequence; corn; ss.	
XX		
OS	Zea mays subsp. mays.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128234P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	19-APR-1999; 99US-0130077P.	
PR	21-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	06-MAY-1999; 99US-0132487P.	
PR	07-MAY-1999; 99US-0132863P.	
PR	11-MAY-1999; 99US-0134256P.	
PR	14-MAY-1999; 99US-0134218P.	
PR	14-MAY-1999; 99US-0134219P.	
PR	14-MAY-1999; 99US-0134221P.	
PR	14-MAY-1999; 99US-0134370P.	
PR	18-MAY-1999; 99US-0134768P.	

PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-01351124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 18-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.

Qy 682 AAGTGATCTCCACAAGAGCTTTGATGATAGATTGAAACATTCAAAAACGTTCAAGGA 741
 |||||
 Db 693 AACATTTATCTACAAGATCGTATGATGATAGATTACAGACTCTTCAGCATGTCGGTGA 752
 |||||
 Qy 742 TCTGATTAAGGCGATGACAGAGTGGTATTCTTGGTCTTGGTGAGACCAAGAGACCGT 801
 |||||
 Db 753 GCTGGAATAAGCGTCTGCTCAGGTGGAATATTGGTCTTGGAGGCGGAGGAAGACCGT 812
 |||||
 Qy 802 GTATCTTTCTCTACACCTTGGCCACCAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
 |||||
 Db 813 GTAGGCTGTTGCATACACTGGCCACTTGGCCACACACACCCAGAGTGTTCCTATCAT 872
 |||||
 Qy 862 AGACTGTGCCAATCAAGGGCACGCCCAATGATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921
 |||||
 Db 873 GCATTGATTGCTGTCAAAGGCAAGCCCTCTTCAGG-----ATCAGAACCTGTAGAG 923
 |||||
 Qy 922 GTTATGATGATTGACAGAACCAATTGCTACTGCAAGATTGGTCATGCCAAGCTATTATTC 981
 |||||
 Db 924 ATATGGGAATGATCCGATGATTGCCAGCGCTCGGATTGTGATGCCAAAGGCAATGGTG 983
 |||||
 Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGCGAGAAAGAGTGATGCTTCATGGCT 1041
 |||||
 Db 984 AGCTTTGGCAGGGCGAGTACGGTTCTCCATGCGAGAGCAAGTCTCTGCTTCTTGGCT 1043
 |||||
 Qy 1042 GGTGTAATGCCATCTTTCACAGGTAAGAAAATGCTCACAACAATGTGTAAAGCGTGGGAT 1101
 |||||
 Db 1044 GGGCCCAACTCGATCTTTCGCCGGTGAAGAAGCTCTGACACTGCAAAACAGGACTTTTGT 1103
 |||||
 Qy 1102 GAGGATAAGCCATGTTGGCTAAATGGGGTCTGAACCAATGG 1144
 |||||
 Db 1104 GCGGACCAGGCAATGTTCAAGATCCTTTGGCCTGATTCCTCCAGG 1146
 |||||

RESULT 15

ABR01975
 ID ABA01975 standard; cDNA; 1439 BP.
 XX
 AC ABA01975;
 DT 06-FEB-2002 (first entry)
 XX
 DE Maize biotin synthase coding sequence #6.
 XX
 KW Maize; biotin synthase; biotin biosynthesis; herbicide;
 KW plant development; clone cca.pk0012.g11.f15; transgenic plant; ss.
 XX
 OS Zea mays.

Key Location/Qualifiers
 FH 115..1248
 CDS /*tag= a
 FT /product= "biotin synthase"
 FT

XX US2001039042-A1.
 PN

XX 08-NOV-2001.
 PD

XX 19-DEC-2000; 2000US-00740288.
 PF

XX 21-DEC-1999; 99US-0172929P.
 PR

XX (ALLEN/) ALLEN S M.
 PA

PA (KINN/) KINNEY A J.
 PA

PA (MIAO/) MIAO G.
 PA

PA (OROZ/) OROZCO E M.
 PA

XX Allen SM, Kinney AJ, Miao G, Orozco EM;
 PI

XX WPI: 2002-040723/05.
 DR

XX P-PSDB; AAM51982.
 DR

XX New polypeptides, useful as targets for herbicide discovery, and as
 PT

PT probes for genetic and physical mapping of genes of which they are part,
 PT or creating transgenic plants, comprises biotin synthase polypeptides and
 PT encoding polynucleotides.

PS Claim 5; Page 29-30; 46pp; English.

XX The present invention provides the protein and coding sequences of biotin
 CC synthase enzymes from barley, wheat, maize, soybean and the prickly
 CC poppy. These sequences can be used to produce transgenic plants which
 CC express different levels of the gene, or express it at different times in
 CC plant development. They can also be used as a target in the production of
 CC herbicides. The present sequence is a maize biotin synthase cDNA obtained
 CC from clone cca.pk0012.g11.f15

XX SQ Sequence 1439 BP; 396 A; 364 C; 365 G; 314 T; 0 U; 0 Other;

Query Match 26.1%; Score 310.2; DB 6; Length 1439;

Best Local Similarity 59.0%; Pred. No. 5.7e-81;

Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTTGACCAAGAGAAATTAAGCTATATATGACACACACCTCATGGACTTGATG 261
 |||||
 Db 235 AACGACTGGAGCGCGCCGAGATCCAGCCGCTTACGACTCAGCGCTCTCGACTCCTC 294
 |||||
 Qy 262 CACTATGCTCAGGTGCAACACAGAAAGTTCCAAACCTTTCAGAGTTCAATGTGCACT 321
 |||||
 Db 295 TTTTCAGGGGCTCAGGTCCACAGAAATGTCATAAATTCAGAGAAAGTGCAGCAATGCAC 354
 |||||
 Qy 322 CTTATGATATCAAACTGGTGTGTACCGAGACTGTAACTACTGTGCCAATCAG 381
 |||||
 Db 355 CTTCTTTCAATCAAGACTGGTGTGATGCGTGAAGATTGTTCTTACTGTCTCAGTCA 414
 |||||
 Qy 382 CGTTACAACTGCTGTCAAGGCTGAAAGAAATCATCAAGTTGATGAGGTGATTGAAGCT 441
 |||||
 Db 415 AGATACAACTGATGATGAAGGCCAAATTCATGAACAAATATGCTGTCTTGGAGCA 474
 |||||
 Qy 442 GCAAGGAGCAAGGCAATGGATCTACAAAGTTCTGTATGGTGTCTGTGGAGAGAG 501
 |||||
 Db 475 GCAAAAAGGCAAGAGTCTGGGAGCACCCGTTTTTGTGATGGAGGCTGCATGGAGAA 534
 |||||
 Qy 502 ATCAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGCTCATGAC 561
 |||||
 Db 535 ACCATTGGCAGGAAATCAAACTTCAACCCAGATTCTTGAATATGTCAGGAAATGAAGGG 594
 |||||
 Qy 562 CTTGGAATGGAGAGTTGTGTCACCCCTGGGAATGGTTGATAAAGCAACCAAGCCTGAA 621
 |||||
 Db 595 ATGGGATGGAGTCTGTTGCCACTAGGCATGATAGAGAAACAAAGCTGGAAGACTC 654
 |||||
 Qy 622 AAAAGTGTGGTGAACGGGTACAAACCAATGATACTTTACAGGAACACTATCCA 681
 |||||
 Db 655 AGAAGGCTGGACTTACAGCATATAATCAACCTAGATACATCAAGAGAGTATTATCCC 714
 |||||
 Qy 682 AGGTGATCTCCACAAGAGCTTTGATGATAGATTGAAAACATTCAAAACGTTCAAGGA 741
 |||||
 Db 715 AACATTATTACCAAGATCATATGATGATAGACTGCACTCTTGAGCATGTCCTGGTGA 774
 |||||
 Qy 742 TCTGGAATTAAGGCGATGACAGGTGTTCTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801
 |||||
 Db 775 GCTGGAATGAAGCATCTGCTCAGGTGAATCAATTTGGTCTTGGTGAGAGGAGGACCG 834
 |||||
 Qy 802 GTATCTTTCTCTACACCTTGGCCACAATGGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
 |||||
 Db 835 GTAGGTTGTTGCATACCCCTAGCTACCTTGGCCTTACACACCCAGAGAGCGTTCCTATT 894
 |||||
 Qy 862 AGACTGTCCTCAATCAAGGCGACGCCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921
 |||||
 Db 895 GCATTGGTTGCTGTAAAGGCGACACCTCTTGGG-----ACCAGAAGCGCTGTAGAG 945
 |||||
 Qy 922 GTTATGATGATTGTCAAGAACCAATGCTACTGCAAGATTGGTGTGCAAGCTCTATTATC 981
 |||||
 Db 946 ATCTGGAAATGATCCGCATGATCGCCACTGCTCGGATCAGATGCCAAGGCAATGGTG 1005
 |||||
 Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGCAAGAGTGTGCTTCTCATGGCT 1041
 |||||

Db	1006	AGGCTTTACGACGCCGAGTACGGTTCTCGATGCCAGAACAAAGCGCTGTGCTTCTCGCT	1065
Qy	1042	GGTTGTAATGCCATCTTCACAGGTAAAGAAATGCTCACAACAATGTGTAAACGGCTGGAT	1101
Db	1066	GGGGCCAACTCCATCTTTGCCGGGAGAACTTCTCACAACCCGCAACACGACTTTGAT	1125
Qy	1102	GAGGATAAAGCCATGTTGGCTAAATGGGGTCTGAAACCAATGG	1144
Db	1126	GCGGACCGCGATGTTCAAGATCCTTGGCTGATCCCCAAGG	1168

Search completed: September 17, 2005, 17:54:06
Job time : 737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 16:11:49 ; Search time 236 Seconds
(without alignments)
8236.851 Million cell updates/sec

Title: US-09-752-957F-1

Perfect score: 1188

Sequence: 1 atgcgtttatattgactgc.....agggtgcattcggtgttga 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:**

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414.4	34.9	1263	4	US-09-248-796A-4445
2	328.2	27.6	3083	1	US-08-346-849-1
3	328.2	27.6	3083	2	US-08-293-284A-1
4	328.2	27.6	3083	4	US-08-898-300-1
5	328.2	27.6	3083	4	US-08-824-513-1
6	304	25.6	1351	2	US-08-401-068-13
7	304	25.6	1351	2	US-08-846-338-13
8	236	19.9	1017	4	US-09-328-352-2698
9	230.4	19.4	954	4	US-09-902-540-7391
10	230.4	19.4	4183	4	US-09-902-540-709
11	211.6	17.8	1830121	4	US-09-557-884-1
12	211.6	17.8	1830121	4	US-09-643-990A-1
13	208.2	17.5	1230025	4	US-09-198-452A-1
14	208.2	17.5	1230230	4	US-09-438-185A-1
15	205.2	17.3	1420	3	US-09-180-109A-23
16	203.6	17.1	1336	3	US-09-180-109A-28
17	200	16.8	640681	4	US-09-790-988-1
18	198.8	16.7	1053	4	US-09-543-681A-1175
19	195.2	16.4	1182	4	US-09-252-991A-3220
20	195.2	16.4	3150	4	US-09-252-991A-3348
21	193.8	16.3	1041	2	US-08-401-068-7
22	193.8	16.3	1041	2	US-08-846-338-7
23	192.4	16.2	5872	3	US-08-411-768B-1
24	192.4	16.2	5872	3	US-08-411-768B-6
25	191	16.1	1358	3	US-09-180-109A-25
26	186.4	15.7	3501	4	US-09-809-665A-70
27	184.2	15.5	34001	4	US-09-596-002-18

ALIGNMENTS

RESULT 1

US-09-248-796A-4445

; Sequence 4445, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICP

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 4445

; LENGTH: 1263

; TYPE: DNA

; ORGANISM: Candida albicans

; US-09-248-796A-4445

Query Match 34.9%; Score 414.4; DB 4; Length 1263;

Best Local Similarity 66.8%; Pred. No. 3.4e-122;

Matches 643; Conservative 0; Mismatches 301; Indels 18; Gaps 3;

Qy	207	CTGGACCAAGAAATTAAAGCTATATATGACACACCACTCATGGACTTGATGCACTA	266
Db	297	CTGGACCAAGATGAATTAAGCAAAATTTACATGCTCCATGATGATTTGATTTCAA	356
Qy	267	TGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTTGSCATCTTAT	326
Db	357	CGTCAATTGAAACATCGTGAGTACCAAAACCTCTGGAAGTCCAATTATGACTCTGT	416
Qy	327	GAATATCAAACTGGTGGTTGTACCGAGGACTGTAAGTACTTGCCCAATCACAGGTTA	386
Db	417	TAACATTAAGCAGGGTGGATGTACCGAAACCTGTTCTTACTGTTTCAATCATCGAACA	476
Qy	387	CAACACTGGTCAAGGCTGAAGATCATCCAAAGTTGATGAGTGTGATTAAGTCGAA	446
Db	477	TGATATCTGAGGTTCAAGCTGAAAAA---CTTGACTGGATGCTGTGATGACATTTGCTAG	533
Qy	447	GGAGGCAAGGGCAATGGATCTCAACAGGTTCTGATGGGTGCTGCTTGGAGAGATGAA	506
Db	534	AGATGCAAGCAGGAGTGGTACTAGAGTTTGCCTTGGTTCCGCTTGGAGATATGCA	593
Qy	507	AGTGAAGAGTCAACTTGAAAGAAATCAAAGAGATGATCCTGCTGTGTCATGAC---CT	563
Db	594	AGGAAGAAAGTCTGCCTTAGACAAAATCAGTAAAGTTGTTAAACAAATCAACGACGATT	653

LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1292..2590
OTHER INFORMATION: /product= "zuotin"
US-08-898-300-1

Query Match 27.6%; Score 328.2; DB 4; Length 3083;
Best Local Similarity 72.1%; Pred. No. 2.9e-94;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

Qy 553 GTCCATGACCTTGGATGAGAGTGTGTCACCCCTGGGAATGGTTGATAAAGACCAAGCC 612
Db 6 GTGAATGATATGGGGCTAGAACGTTGTACTTTAGGTATGGTTGATCAAGATCAAGCA 65

Qy 613 ACTGAATTGAAAAGTCTGGGTTGACGGCTGACACCAATACATTAATGATCTTCAAGGAA 672
Db 66 AAGCAATTAAGATGCGAGTTTGAATGATACCAACCAATACATCGACATTCAGAGAA 125

Qy 673 CACTATCCAAAGGTGATCTCCACAAAGAGCTTTGATGATAGATTGAATAAATTCAAAAC 732
Db 126 CACTATAGTAGTTCATCCACGAGAACCTACGACGACAGGTTACAGACCATCAAGAA 185

Qy 733 GTTCAAGCATCTGGATTAAAGCATGACAGGTGTATCTTGGTCTTGGTGAGACCAA 792
Db 186 GTCCAAGAAATCTGGAAATAAAGCCGTATACCGGTGTATTTGGGTCTCGGTGAAGCGAA 245

Qy 793 GAGGACCGTGTATCTTCTCTACACCTTGGCCCAATGGATCAGATCCAGAGTCTCTT 852
Db 246 GACGACCATATAGGATTCATCTACACATTAATCAATATGTCTCTCATCTGAGTCCCTA 305

Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGGACGCCAATGTATGAAGAAGTT----- 903
Db 306 CCAATTAATAGACTAGTTGCTATCAAGGGAGCTCCAATGGCTGAGGAACCTTGGCGATCCA 365

Qy 904 AAGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCATGCTACTGCAAGATTGGTC 963
Db 366 AAGAGTAAAAAGTTGCAATTCGAGAAATTTGAGAACCATTTGCCACAGCGAGAGATGTT 425

Qy 964 ATGCCAAGCTTATATCAGATTGGCTGCAGGAAGATATACAATGAAGAGCGAGAA 1023
Db 426 ATGCCAAGGCCATTAAGACTTGGCCCTGGTCTGTATACATGAAGAAGAGAGGAA 485

Qy 1024 GTGATGCTTCAAGCTGGTGTGTAATGCCATCTTTCACAGGTAAAGAAATGCTCACAA 1083
Db 486 TTTGTCTCTTTCATGGCAGGTTGTAACAGTATCTTCCCGGTAAAGAAATGCTGACGACA 545

Qy 1084 ATGCTAACGCTGGGATGAGATTAAGCCATGTTGGCTAATGGGTCTGAAACCAATG 1143
Db 546 ATATATAACGTTGGGAGAGAGAAAGGCAATGTTGGCTAAATGGGATTCGAACCTATG 605

Qy 1144 CAGAGTTTCAAAATACAA 1160
Db 606 GAGGCATTTAAGTACGA 622

RESULT 5

US-08-824-513-1
Sequence 1, Application US/08824513
Patent No. 6800481
GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuang
APPLICANT: Rich, Alexander
APPLICANT: DiPersio, C. Michael
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR

NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,513
FILING DATE: March 26, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/293,284
FILING DATE: August 22, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1292..2590
OTHER INFORMATION: /product= "zuotin"
US-08-824-513-1

Query Match 27.6%; Score 328.2; DB 4; Length 3083;
Best Local Similarity 72.1%; Pred. No. 2.9e-94;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

Qy 553 GTCCATGACCTTGGATGAGAGTGTGTCACCCCTGGGAATGGTTGATAAAGACCAAGCC 612
Db 6 GTGAATGATATGGGGCTAGAACGTTGTACTTTAGGTATGGTTGATCAAGATCAAGCA 65

Qy 613 ACTGAATTGAAAAGTCTGGGTTGACGGCTGACACCAATACATTAATGATCTTCAAGGAA 672
Db 66 AAGCAATTAAGATGCGAGTTTGAATGATACCAACCAATACATCGACATTCAGAGAA 125

Qy 673 CACTATCCAAAGGTGATCTCCACAAAGAGCTTTGATGATAGATTGAATAAATTCAAAAC 732
Db 126 CACTATAGTAGTTCATCCACGAGAACCTACGACGACAGGTTACAGACCATCAAGAA 185

Qy 733 GTTCAAGCATCTGGATTAAAGGCATGACAGGTGGTATCTTGGTCTTGGTGAGACCAA 792
Db 186 GTCCAAGAAATCTGGAAATAAAGCCGTATACCGGTGTATTTGGGTCTCGGTGAAGCGAA 245

Qy 793 GAGGACCGTGTATCTTCTCTACACCTTGGCCCAATGGATCAGATCCAGAGTCTCTT 852
Db 246 GACGACCATATAGGATTCATCTACACATTAATCAATATGTCTCTCATCTGAGTCCCTA 305

Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGGACGCCAATGTATGAAGAAGTT----- 903
Db 306 CCAATTAATAGACTAGTTGCTATCAAGGGAGCTCCAATGGCTGAGGAACCTTGGCGATCCA 365

Qy 904 AAGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCATGCTACTGCAAGATTGGTC 963
Db 366 AAGAGTAAAAAGTTGCAATTCGAGAAATTTGAGAACCATTTGCCACAGCGAGAGATGTT 425

Qy 964 ATGCCAAGCTTATATCAGATTGGCTGCAGGAAGATATACAATGAAGAGCGAGAA 1023
Db 426 ATGCCAAGGCCATTAAGACTTGGCCCTGGTCTGTATACATGAAGAAGAGAGGAA 485

Qy 1024 GTGATGCTTCAAGCTGGTGTGTAATGCCATCTTTCACAGGTAAAGAAATGCTCACAA 1083
Db 486 TTTGTCTCTTTCATGGCAGGTTGTAACAGTATCTTCCCGGTAAAGAAATGCTGACGACA 545

Qy 1084 ATGCTAACGCTGGGATGAGATTAAGCCATGTTGGCTAATGGGTCTGAAACCAATG 1143
Db 546 ATATATAACGTTGGGAGAGAGAAAGGCAATGTTGGCTAAATGGGATTCGAACCTATG 605

Qy 1144 CAGAGTTTCAAAATACAA 1160
Db 606 GAGGCATTTAAGTACGA 622

Db 366 AAGAGTAAAGTTGCAATTCGACGAATTTTGAAGACCAATTCGCCACAGCGAGAAATAGTT 425
Qy 964 ATGCCAACGCTTATATCAGATTGGCTGCAGGAAGATATACAATGAAGAGCGAGAACAG 1023
Db 426 ATGCCAAAGGCCATTATAAGACTTGGCGTGTATACAAATGAAGAAACAGAGCAA 485
Qy 1024 GTGATGCTTCATGGCTGGTTGTAATGCCATCTTCACAGTAAGAAAATGCTCACAA 1083
Db 486 TTGTCTGTTTCATGGCGAGTTGTAACAGATATCTTCACCGTAAGAAAATGCTGACGCA 545
Qy 1084 ATGTGTAACGCTGGGATGAGGATAAGCCATGTTGGCTAAATGGGCTCTGAAACCAATG 1143
Db 546 ATATATACGCTGGGACGAAGACGAATGTTGGCTAAATGGGATTCGAACCTATG 605
Qy 1144 GAGAGTTTCAATAACAA 1160
Db 606 GAGGCATTTAAGTAGCA 622

RESULT 6

US-08-401-068-13
; Sequence 13, Application US/08401068
; Patent No. 5859335
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Enhanced Biotin Biosynthesis in Plant Tissue
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,068
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/351,970
; FILING DATE: 08-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 59..1192
; OTHER INFORMATION: /product= "Arabidopsis BioB enzyme"

US-08-401-068-13

Query Match 25.6%; Score 304; DB 2; Length 1351;
Best Local Similarity 58.6%; Pred. No. 1e-86;
Matches 551; Conservative 0; Mismatches 380; Indels 9; Gaps 1;
Qy 202 AACACCTGGACCAAGAGAAAATTAAGCTATATATGACACACACTCATGGACTTGATG 261
Db 194 AACGATTGGAGTAGAGATGAATCAAGTCTGTGTTATGATTCTCTCTTGTGACCTCCTC 253

Qy 262 CACTATGCTCAGGTGCAACACACAGAGGTTCCAAACAACCTTCAGAGGTTCAATTGTGCACT 321
Db 254 TTCATGGAGCTCAGGTTTCATAGACATGTTTCATAAATTCAGGGAGGTACAACAATGTACC 313
Qy 322 CTTATGATATCAAACTGGTGTGTACCGAGGAGCTAAGTACTGTGCCCAATCACAG 381
Db 314 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGACTGTTTCATATTGTCTCAGTCTTCG 373
Qy 382 CGTTACAACTCGTGTCAAGGCTGAAAGAAATCATCCAAGTGTATGAGGTGATTGAAGCT 441
Db 374 AGATATAGCACTGGAGTTAAGGCACAAAGACTCATGTCTAAGGACGCTGTCATTGTGCT 433
Qy 442 GCAAGAGGACAAAGGCCAATGATCTACAAGGTTCTGTATGGTGTCTGTGGAGAGAG 501
Db 434 GCTAAGAAAGGCAAAAGAGCTGGGAGCACACGTTTTTGGCATGGTGTCTGTGGCGAGAT 493
Qy 502 ATGAAAGGTAGAAAGTCAAACTTTGAAGAAATCAAGAGATGATCACTGCTGTCATGAC 561
Db 494 ACAATTGGACGGAACCAACTTCAGCCAGATTTCTTGAATACATCAAGAAATTAAGAGGC 553
Qy 562 CTTTGGATGGAGAGTTGTGTCACTGGGAATGGTTGATAAAGACCAAGCCACTGAATTG 621
Db 554 ATGGGATGGAGTTTGTCTGCACTTAGGCAATGTTGAGAAACAAAGCACTAGAGCTA 613
Qy 622 AAAAGTCTGGTGTGACGGCGTACAAACATTAACATTGATCTTACAAGGAACACTATCCA 681
Db 614 AAGAAGCTGGCTCTACTGCTTATAACCAACAATCTTGTATCTTCAAGAGAGTACTACCCA 673
Qy 682 AAGGTATCTCCACAAGAGCTTTGATGATAGATTGAAAACATTCAAAAAGCTTCAAGGA 741
Db 674 AACGTCACTACTAGAAAGTTATGACGATCGCCTTGAAACTCTTAGCCATGTTCTGAT 733
Qy 742 TCTGGATTAAAGGCATGCACAGGTGTATTCTTGGTCTTGGTGCAGACCCCAAGAGGACCGT 801
Db 734 GCTGGAATCAACGTTTGTTCAGGAGGAATCATTAGGCTTGTGAGGAGAGGAGACAGA 793
Qy 802 GTATCTTTCTCTACACCTTGGCCACAATGGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
Db 794 ATAGGTTTATTACACACGCTGGCAACACTTCTCTCACCTCAGAGTGTTCCTCATTAAT 853
Qy 862 AGACTGCTCCCAATCAAGGCGACGCCAATGTATGAAGAGTTAAGAAACAGAGGTTGAA 921
Db 854 GCTCTACTTGCAGTGAAGGCACTCTCTTTGAAG-----ACCAGAAAGCCAGTTGAG 904
Qy 922 GTTGTAGAGATTGTCAAGAACCTTGTCTACTGCAAGATTGTCATGCCAAGCTCTATTATC 981
Db 905 ATATGGAGATGATCAGGATGATTGGAACCGCAGTAATTGTAATGCCAAAGGATGGTG 964
Qy 982 AGATTGGCTGCAGGAAGATATACAATGAAGAGGCGAGAACAGGTGATGCTTCTATGGCT 1041
Db 965 AGACTGCTGCTGGTAGAGTCCGGTTCTCAATGTCCGACACAGCTCTCTGTTTCTTGTCT 1024
Qy 1042 GGTGTAATGCCATCTTCACAGTAAGAAAATGCTCACAAATATGTGAACGGCTGGAT 1101
Db 1025 GGTGCAAACTCTATCTTCCACCGAGAGAGCTTTTAACCAACACCAACAATGATTTTAC 1084
Qy 1102 GAGGATAAGCCATGTTGGCTAAATGGGCTCTGAAACCAA 1141
Db 1085 GCTGACCACTCATGTTCAAGACATTAGGCTCAATTCCTA 1124

RESULT 7

US-08-401-068-13
; Sequence 13, Application US/08401068
; Patent No. 5869719
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Transgenic Plants Having Increased Biotin Content
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5869719artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005

CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,338
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 59..1192
OTHER INFORMATION: /product= "Arabidopsis biotin synthase enzyme"
US-08-846-338-13

Query Match 25.6%; Score 304; DB 2; Length 1351;
Best Local Similarity 58.8%; Pred. No. 1e-86;
Matches 551; Conservative 0; Mismatches 380; Indels 9; Gaps 1;

QY 202 AACCTGACCAAGAAATTAAGCTATATATGACACCACTCATGACCTTGATG 261
DB 194 AACGATTGAGTAGAGATGAATCAAGTCGTGTATGATCTCTTCTGACCTCTC 253
QY 262 CACTATGCTCAGGTGCAACACAGAAAGTTCCAAACCTTCAGAGTTCAATTTGTCAT 321
DB 254 TTCCATGGAGCTCAGGTTTCATGACATGTTTCAAACTTCAGGAGGTACAACAATGTACC 313
QY 322 CTTATGAATATCAAACTGGTGTACCGAGGACTGTAAAGTCTGTATGTTGCTCCCAATCACAG 381
DB 314 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGACTGTTTCATATGTTCTCAGTCTTCG 373
QY 382 CGTTACCAACACTGGTGTCAAGGCTGAAGAAATCATCAAGTTGATGAGGTGATTGAAGCT 441
DB 374 AGATATAGCACTGGAGTTAAGGCAAAAGACTCATGTCCTAAGACGCTGTCAATGATGCT 433
QY 442 GCAAGAGGCAAGGCCAATGGATCTCAAGGTTCTGTATGGTGTCTGCTTGGAGAGAG 501
DB 434 GCTAAGAGGCAAGAAAGCTGGAGCACACGTTTGTGATGGTGTCTGCTTGGCGAGAT 493
QY 502 ATGAAGGTAGAAAGTCAAACTTGAGAAATCAAAAGAGATGATCACTGCTGCCATGAC 561
DB 494 ACAATTGGACGGAAGAAACCAACTTCAGCCAGATTCTTGAATATACATCAAGAAATAAGAGGC 553
QY 562 CTTGGAATGGAGATTGTGTCACTTGGGAATGGTTGATAAAGACCAAGCCACTGAATTG 621
DB 554 ATGGGAGTGAAGTTGTCTGACCTTAGGATGATGTTGAGAAACCAAGCACTAGAGCTA 613
QY 622 AAAAGTGTGGTTGACGGCGTACAAACATGATGATGATGATGATGATGATGATGATGATGAT 681
DB 614 AAGAAGGCTGGCTCACTGCTTATAACCAAACTTGTGATCTTCAAGAGAGTACTACCCA 673
QY 682 AAGTGATCTCCACAGAGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
DB 674 AACGTCATCACTACTAGAAAGTTATGACGATCGCTTGAAGAACTTCTAGCCATGTTGCTGAT 733

QY 742 TCTGGATTAAAGGCATGACAGGTGGTATTCTTGTGTTGTGAGACCAAGAGGACCT 801
DB 734 GCTGGAATCAACGTTTGTTCAGGAGGAATCATAGGCTTGGTGAGCAGAGGAAGACAGA 793
QY 802 GTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTCCCAATCAAC 861
DB 794 ATAGGTTTATTACACACGCTGGCAACACTTCTTCTCACCTCGAGAGTGTCCCATTAAT 853
QY 862 AGACTGGTCCCAATCAAGGCGACGCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921
DB 854 GCTCTACTTGTGAGTGAAGGCACTCTCTTGAAG-----ACCAGAGCCAGTTGAG 904
QY 922 GTTGATGAGATTGTGAGAACCAATCTCTACTGCAAGATTGGTCAATGCCAACGCTTATATC 981
DB 905 ATATGGGAGATGATCAGGATGATTGGAACCGCACCTGATGTATGTAATGCCAAAAGCATGGT 964
QY 982 AGATTGGCTGACGGAAGATATACATGAAGAGGCGAGACAGGTGATGCTGCTTCATGGCT 1041
DB 965 AGACTGTCTGTGGTAGAGTCCGGTCTCTCAATGTCCGAACAAGCTCTCTGTTCCTTGTCT 1024
QY 1042 GGTGTAATGTCATCTTCAAGGTAAGAAATGCTCACAACAATGTGTAACGGCTGGGAT 1101
DB 1025 GGTGCAAACTCTATCTTCAACGAGAGAAAGCTTTTAAACACACCAACAATGATTTGAC 1084
QY 1102 GAGATAAAGCCATGTTGGCTAAATGGGCTCGAAACCAA 1141
DB 1085 GCTGACGAGTCATGTTCAAGACATTAGGCTTCATTCCTA 1124

RESULT 8
US-09-328-352-2698
Sequence 2698, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 2698
LENGTH: 1017
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-2698

Query Match 19.9%; Score 236; DB 4; Length 1017;
Best Local Similarity 54.7%; Pred. No. 5.9e-65;
Matches 520; Conservative 0; Mismatches 415; Indels 15; Gaps 2;

QY 130 AAGACACGAGTCAACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTC 249
DB 28 ATGACTCTACGTAATGATGAGTCTCGTGAAGAAATCCAAAGCTCTATATGAACAACCTTTT 87
QY 250 ATGGACTTGTGACACTATGCTCAGGTGCAACACAGAGGTTTCCAAACACCTTCAGAGGTT 309
DB 88 TTGGATTTAGTTTTTAAGGCTCAACAGTACACCGTGAGCACTTCACTGCCAATACAATT 147
QY 310 CAATTGTGACCTCTTATGAATATCAAACTGGTGGTGTGACGAGAGCTGTAACTACTGT 369
DB 148 CAGGTGACGACCCCTTTTATCTATTAAACGGGTAAATGCTCTGGAAGATTGCAAAATATTGC 207
QY 370 GCCCAATCACAGCTTACACACTGCTGTCAGGCTGAAAGAAATCATCCAGTTCTGATGAG 429
DB 208 TCTCAATCTGCACATTACGATTCAAAATTAGAAGCAGAAACGTAATGCTGTTGAAANA 267
QY 430 GTGATTGAAGTGGCAAGGAGGCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 489
DB 268 GTAATTAGCGAAGCAAAAGCTGCAAAAGATTTCAGGTTTCATCTCGTTTTTGTATGGTGTCT 327
QY 490 GCTTGGAGAGATGAAGGTTGAAGAAAGTCAAACTTGAAGAAATCAAAAGAGATGATCACT 549

Db 328 CTTTGGCG-----CAACCCCTCATGAGCGGATATGCTTATGTTTTAGAAATGGTACGT 381
Qy 550 GCTGTCATGACCTTGAATGAGAGTGTGTCAACCTCGGAAATGGTTGATAAGACCAA 609
Db 382 GAAGTTAAGCATTAGGTATGGAACCTGTATGACTTTAGTATGCTTATCATCAACAA 441
Qy 610 GCCACTGAATTGAAGTGTGGTTGACCGCGTACCAACCAATATGATGATCAAG 669
Db 442 GCAGAACGCTTAAAGATGCGGTTTAGACTATTACCAACCAATACCTTGATCTCTCGC 501
Qy 670 GAACACTATCAAGGTGATCTCCACAAGAGCTTTGATGATAGATGAAACATTCAAA 729
Db 502 GAGTATTAATCCCAATCATCAGACACAGTACTTTTATGATGACCGTTTGAATACTTTAGAT 561
Qy 730 AACCTTCAAGGATCTGGATTAAAGGCATGACAGGTGATTTCTTGGTCTTGGTGAGACC 789
Db 562 TATGTAGCTCAAGCTGGTATGAAAGTATGATGTCGCGTATTTGGTTTAGGTGAGAGC 621
Qy 790 CAAGAGGACCGTGTATCTTCTCTACACCTTGGCCCAATGATGATCAGCATCCAGATCT 849
Db 622 CGTGAAGATCGCATTGGCTTGTATACATGAGTTAGCAACCTTACCTATTATCCAGATCT 681
Qy 850 CTTCCAAATCAAGACTTGGTCCCAATCAAGGGCAGCGCAATGATGATGAAGAGTTTAAGAC 909
Db 682 GTGCCAATCAACATGCTTGTTCCAATTCAGAGTACACCAATTAGCTGATGTTG----- 733
Qy 910 AAGCAAGTTGAGTGTGATGATGTCAGAACCAATGCTACTGCAAGATTGGTCATGCCA 969
Db 734 -AAACTAGATGTTATGAAAGTATGATGTCGTAATTCGCGTACGATCATTTATGCCG 792
Qy 970 ACGTCTATTATCAGATTGGCTGCGAGGAGATATACAAATGAAGAGGCGGACAGGTCATG 1029
Db 793 CATAGTATATTCGTTTATCTGCTGGTCTGTAATCTTTAAGTACTCAGATCAAGCTTTA 852
Qy 1030 TGCTTCATGCTGTTGTAATGCCATCTTCAAGGTGAAGAAATGCTCACAAATGTGT 1089
Db 853 GCATTTATGGCGGTGCAAAATTCCTGTTCTGTGTGATAAATTTACTTACGAGCCCAAT 912
Qy 1090 AACGGCTGGATGAGGATAAGCCATGTTGGCTAAATGGGCTCGAAACC 1139
Db 913 GCAGGTGAAGTTAAGACCAAGCGTTATTTAATAAATTAGGTTAAGC 962

RESULT 9

US-09-902-540-7391
; Sequence 7391, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7391
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7391

Query Match 19.4%; Score 230.4; DB 4; Length 954;
Best Local Similarity 54.5%; Pred. No. 3.6e-63;
Matches 512; Conservative 0; Mismatches 416; Indels 12; Gaps 2;
Qy 207 CTGGACCAAGAGAATTAAGCTATATATGACACACCACTCATGACTTGTATGACTA 266
Db 9 CTGGTCGCTGTCGAGGTGAGGCGCTCTACGAACCTCCGCTGCTGACACCTGTCGACAA 68

Qy 267 TGCTCAGGTCAACACACAGAAAGTTCCAAACAACTTCAGAGGTTCAATTTGTCATCTTAT 326
Db 69 GGCCCAAGCGTCCACCGGCCGTGTTGCTGGAACAACAGAGTCCAGCTCTGTTGCTGCT 128
Qy 327 GAATATCAAAACTGGTGGTTGATCCGAGGAGTGTAACTGATGTCGCCCAATACACAGCGTTA 386
Db 129 GTCCATCAAGACCGGCGGCTGCCCCGAGGAGTGTGCTGATCTGCCCGCAGCGGCTCGCTA 188
Qy 387 CAACACTGGTGTCAAGGCTCAAAAGAAATCATCCAAGTTGATGAGGTGATTAAGCTGCAAA 446
Db 189 CARGACGGGCTCAAGCGGAGAGCTGATGGCGTCCCGAGCTGCTGATGATGCGCGTC 248
Qy 447 GSAGGCAAGGCCCAATGGATCTTCAAGGTTCTGTATGGGTGCTGCTTGGAGAGATGAA 506
Db 249 CAAGGCCCGCGCCGCGAGCCACCGCTTCTGATGGCGCGCGTGGCGGAGGTGAA 308
Qy 507 AGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGACCTTGG 566
Db 309 GSACGG---TCCGCAAGTTCCAGACGCTGCTGGAGATGCTGCGCGCTGCGCGCTGGG 365
Qy 567 AATGGAGAGTTGTGTCACCTCGGAATGTTGATAAGACCAAGCCACTGAATTTGAAAG 626
Db 366 CATGGAGCGGTGTCACCTCGGCAATGCTTCCGAGAGCAGCGGAGCGCTGCGCGA 425
Qy 627 TGCTGGGTTGACGGCGTACAAACATTAACATTTGATCTTCAAGGAAACACTATCCAAAGGT 686
Db 426 GCGCGCTGTCCGCTGATCAACCAACCTGGACACGTCCGCCGAGCACTACGCGGACAT 485
Qy 687 GATCTCACAAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 746
Db 486 CATCTCCACCGCACCTTATGAAGACCGCTGCGCACGCTCAACCGCTGCGCGACGCGG 545
Qy 747 ATTAAAGGCATGACAGGTGTTTCTGCTGTTGTTGAGACCCCAAGAGCACCGTGTATC 806
Db 546 CATCTCGGTGCTGCGCGCGCATCATTTGSCATGGCGAGTGGTGTGATGACCGCTGCAA 605
Qy 807 TTTCTCTACACCTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATG 866
Db 606 CTTGCTGCGCACCTTGGCCCAACAGAGGACACACCGGAGTGGTGTGCTGCTGCTGCTG 665
Qy 867 GGTCCCAATCAAGGCGACCGCAATGATGATGATGATGATGATGATGATGATGATGATG 926
Db 666 GGTGGCGCTGGAGGCGACCGCTGCGAGG-----AGCAGCAGCGCGTGGAGACGGT 716
Qy 927 TGAGATTGTGAGAACCATTTGCTACTGCAAGATTGTTGATGCCAACGCTCTATTATCAGAT 986
Db 717 GGACATGGTGGCACCATTGCGCACCGCGCATCTGTATGCCCCAGTCCATGGTGGCCT 776
Qy 987 GGCTGAGGAAGATATACAAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1046
Db 777 GTCCGCGCGCGGACAGATGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 836
Qy 1047 TAATGCACTTTCAAGGTGAAGAAATGCTCAACAAATGTTGTAACGCTGGGATGAGGA 1106
Db 837 CAACTCACTCTTCTTCGCGGAGAGCTGCTACCAACCGGCAACCCCGAGTACACCCAGGA 896
Qy 1107 TAAAGGCATGTTGGCTTAATGGGCTGTAACCAATGGAG 1146
Db 897 CATGGCCCTGCTGGAGAGGCGGGAATCCGCCCTTGGAG 936

RESULT 10

US-09-902-540-709/c
; Sequence 709, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 709
; LENGTH: 4183
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-709

Query Match 19.4%; Score 230.4; DB 4; Length 4183;
Best Local Similarity 54.5%; Pred. No. 9.1e-63;
Matches 512; Conservative 0; Mismatches 416; Indels 12; Gaps 2;
QY 207 CTGACCAAGAGAAATTAAGCTATATATGACACCACTCATGACCTGATGACACTA 266
DB 2127 CTGGTCGCTGCGAGGTGAGGCGCTCTAGAACTCCCGCTGCTGACCTGTGTGACAA 2068
QY 267 TGCTCAGGTGCAACACAGAGAGTTCCAAACAACCTTCAGAGGTTCAATTTGCACTCTTAT 326
DB 2067 GGCCAGACGGTCCACCGGCGCTGTTCTGTTGGACAAAGTCCAGCTCTGTTGCTGCT 2008
QY 327 GAATATCAAAACTGGTGTGTACCGAGGACTGTAACTGCTGCCCCAATCAACAGGTTA 386
DB 2007 GTCCATCAAGACGGCGCTGCCCGAGGAGTGTCTGTACTGCCCGCAGCGCGCTGCTA 1948
QY 387 CAACACTGGTGTCAAGCTCAAGAGATCATCCAGTTGATGAGGTGATTTGAAGTCAAA 446
DB 1947 CAAGACGGCGCTCAAGCGGAGAGCTGATGGCGGTGCCGAGCTGCTGGATGCCGCTC 1888
QY 447 GGAGGCAAGGCCAAATGATCTCAAGGTTCTGTATGGGTGCTGCTTGGAGAGAGATGAA 506
DB 1887 CAAGCGCGCGCGCGGAGACCCCGCTTCTGATGGGCGCGCTGGCGGAGGTGAA 1828
QY 507 AGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGACCTGG 566
DB 1827 GGACGG---TCCGAGTTCCACAGCGTCTGGAGATGTTGCGCGCGCTGCGCGCTGGG 1771
QY 567 AATGGAGAGTTGTACCTCGGAGATGTTGATTAAGACCAAGCCACTGAATTTGAAG 626
DB 1770 CATGGAGGCGTGTGCCACCTCGGCATGCTCTCCGAGAGCGAGCGGCGCTGCGCGA 1711
QY 627 TGCTGGTTGACGCGGTACCAACCATTAACATTTGATCTTTACAAGGAACTATCAAAAGT 686
DB 1710 GGCGGCGCTGTCCGCTACCAACCACTGGACAGTCTGCCCGGAGCACTACGCGCAT 1651
QY 687 GATCTCCACAGAGCTTTGATGATGATTAAGAAATCAAAATCAAAAGCTTCAAGGATGG 746
DB 1650 CATCTCCACCGCACTTATGAAGACCGCTGCGCAGCTCAACCGCGTGGCGACGCGG 1591
QY 747 ATTAAGGCATGCACAGGTGTTCTTGGTCTTGGTGAGACCCAGAGACCGTGTATC 806
DB 1590 CATCTCGTGTGTGGCGGCGCATATTTGGCATGGGCGAGTCTGGTATGACCGCTGCAA 1531
QY 807 TTTCTCTACACCTTGCCCAATGATGATGATGATGATGATGATGATGATGATGATGAT 866
DB 1530 CTGCTGCGGACCTGCGCAACAGGAGACACCCCGGAGTCTGGTGGCCATCAACGCT 1471
QY 867 GGTCCCAATCAAGGCGACGCAATGATTAAGAGTTAAGAACAGCAAGTTGAAGTTGA 926
DB 1470 GGTGGCGGTGGAGGCGACGCGCTGACAG-----AGCAGACGCGCTGGAGACGCT 1420
QY 927 TGAGATTGTGACAGCACTTCTACTGCAAGATTGGTTCATGCCAACGCTCTATTATCAGATT 986
DB 1419 GGACATGGTGTGCGACCATCCGCGCGCATCTTGGATGCCCGAGTCCATGTTGGCGCT 1360
QY 987 GGCTGAGGAGATATACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
DB 1359 GTCCGGGCGGCGAGGAGATGAACGAGAGGCGCACTGCTGTGATGATGATGATGATG 1300
QY 1047 TAATGCCATCTTCAAGGTTAAGAAATGCTTCAACAATGTTAAACGGTGGGATGAGGA 1106
DB 1299 CAACCTCACTCTTCTTGGCGGAGAGCTGCTCAACCGGCAACCCCGAGTACACCCAGGA 1240

QY 1107 TAAAGCCATGTTGGCTAAATGGGTCTGAAACCAATGGAG 1146
DB 1239 CATGGCCCTGCTGGAGAGCGGGAATCCGCCCTGGAG 1200

RESULT 11

US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSES: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 17.8%; Score 211.6; DB 4; Length 1830121;
Best Local Similarity 54.3%; Pred. No. 4.7e-55;
Matches 500; Conservative 0; Mismatches 404; Indels 16; Gaps 3;

QY 221 AAATTAAGCTATATATGACACCACTCATGCACTGATGCACTATGCTCAGGTCAAC 280
DB 1085608 AAGTTAGCGGTATTTGAAGCCATTTTGGATTAGTTTACCGAGCAACGCAAGTTC 1085667
QY 281 ACAGAGGTTCCAAACACCTTCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTG 340
DB 1085668 ATCGTAAGCATTTTAATCTCGCGGATTCAGTTATCCACGTTAATGTCTATCAAAACGG 1085727
QY 341 GTGGTTGTACCGAGCACTGTAGTACTGTGCCAATCAGAGGTTACACACTGGTGTCA 400
DB 1085728 GGGGATGCCAGAGATTGTAGTTATTTGCCCTCAATCAGCCCGTTATCATCTGCGGTAC 1085787
QY 401 AGCGTGAAGAAATCATCCAAAGTTGATGAGTGAAGCTGCAAGAGGAGGCAAGGCCA 460
DB 1085788 AAATCAGCAGTTATTAGATGTTGATGAGATCGTCGCTAAAGCAAAATTTGCAAGAC 1085847
QY 461 ATGGATCTCAAGGTTCTGTATGGTGTGCTTGGAGAGAGATGAAGGTGAAGAGTCAAA 520

Db 1085848 GTGTGTCAGGGGTTTCTGTATGGGGCTGTCTGGCGAGGCCCTTAAGCCAA-----AAG 1085901
Qy 521 ACTTGAAGAAATCAAGAGATGATCACTGTCTGCATGACCTTGGATGAGAGTTGTG 580
Db 1085902 ATATTGAGAAAGTACAGAAATATTAAAGCGGTGAATCATCTGTGTTAGAACTTGG 1085961
Qy 581 TCACCTGGGAATGGTTGATTAAGACCAAGCCACTGAATGAAAGTCTGGTTGACGG 640
Db 1085962 GTACCTTTGTTATTGCAAGACGGTATGGCGGAAGATTAAAGAAAGCTGGATTAGATT 1086021
Qy 641 CGTACCAATCAATATTGATCTTACAGGACACTATCCAAAGGTGATCCCAAGAA 700
Db 1086022 ATTACAAATCAATCTTGATAC-CGCAACAGAAATATTGCAAGAGTATGGTACTGCC 1086080
Qy 701 GCTTTGATGATGATTGAAACATCAAAAGCTTCAAGGACTCTGATTAAGGATGCA 760
Db 1086081 GTTTTGTATGTCCTTAGCACCTTAGGAAAGTGGCAAGCTGGATTAAGAGTGT 1086140
Qy 761 CAGTGTGTTATTCTTGGTCTTGGTGAGACCCCAAGAGCCGTGTATCTTTCTTACACT 820
Db 1086141 GTGTGGCATTGTAGGGATGAATGAACCTCGCAAGAACGGCTGGATTAAATCGGAGCC 1086200
Qy 821 TGGCCAAATGGATCAGATCAGAGTCTCTTCCAAATCAACAGACTGTCCCAATCAGG 880
Db 1086201 TTGCAAAATCTTGATCCGCAACCTGAGTCAGTCCGGAATTAATCAACTGTTAAGGTTGAAG 1086260
Qy 881 GCAGCCCAATGATGAAGAGTTAAGAACAGCAAGAGTTGAAGTTGATGAGATTGTCAGAA 940
Db 1086261 GTACGCCCTTAGCCGATCGG-----AGAAATTAGTTGAGCGAGTTTGGCGAA 1086311
Qy 941 CCATTGTCTACTGCAAGATTGGTCTATGCAAGCCCAAGCGTCTATTATCAGATGGCTGCAGGAAGAT 1000
Db 1086312 CGATTGTCTTGGCAGTATTACTATGCCCCAAAAGTTATGTTCCGCTTCTGCAGGTGCA 1086371
Qy 1001 ATACAATGAAGAGCAGCAAGAGTGTGCTTCAATGAGTGTGTTGTAATGCCATCTTCA 1060
Db 1086372 GTGGTATGACTGAAGAAATGCAAGGATGTGTTTATGCGGGTGCAGAAATCTATTTCT 1086431
Qy 1061 CAGGTAAAGAAATGCTCAACAATGTGTACGGCTGGGATGAGGATAAGCCATGTTGG 1120
Db 1086432 ATGGCATTAATTTACTTGTACGATAATCCAGAAAGATGGCGATCATGTTATGATGG 1086491
Qy 1121 CTAATGGGCTGTAACCA 1140
Db 1086492 CAAATTAGATTAGAGCCA 1086511

RESULT 12

US-09-643-990A-1

Sequence 1, Application US/09643990A

Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643.990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PBI86P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match

17.8%; Score 211.6; DB 4; Length 1830121;

Best Local Similarity 54.3%; Pred. No. 4.7e-55;

Matches 500; Conservative 0; Mismatches 404; Indels 16; Gaps 3;

Qy 221 AATAAGCTATATATGACACACCTCATGAGCTATGATGCACTATGCTCAGGTGCAAC 280
Db 1085608 AAGTTGAGGGCTTATTTGAAACGCCATTTTGGAAATAGTTTACGAGCAACGCAAGTTC 1085667
Qy 281 ACAGAAGTTCCAAACACCTTCAGAGTTCAATTTGTCACCTTTATGAATATCAAACTG 340
Db 1085668 ATCTAAGCATTTTAATCTTCGCGCGATTCACTTATCCAGTTATGCTATCAAAACGG 1085727
Qy 341 GTGCTTGTACCGAGACTGTAAGTACTGTGCCCAATCACAGCGTTTACAACACTGGTGCA 400
Db 1085728 GGGGATCCCAAGAGATTGATTGTTTATGCCCTCAATCAGCCCGTTATCATCTGCGGTAC 1085787
Qy 401 AGGTGAAGAAATCATCCAGTTGATGAGGTGATGAAGCTGCAAGAGGAGCAAGGCCA 460
Db 1085788 AAAATCAGCAGTTATTAGATGTTGATGAGATCGTCGCTAAAGCAAAAATTCGAAAGCAC 1085847
Qy 461 ATGGATCTACAAGGTTCTGTATGGTCTGCTTCGAGAGAGATCAAAAGGTAGAAAGTCAA 520
Db 1085848 GTGGTGAGGGCGTTTCTGTATGGGGCTGCTTGGCGAGGCCCTAAGCCAA-----AAG 1085901
Qy 521 ACTTGAAGAAATCAAGAGATGATCACTGTGTCTCATGACCTTTGGAATGGAGAGTTGTG 580
Db 1085902 ATATTGAGAAAGTACAGAAATATTAAAGCGGTGAATCATCTTGGTTTACAACTTGGC 1085961
Qy 581 TCACCTGGGAATGGTTGATTAAGACCAAGCCACTGAATGAAAGTGTGGTTGACGG 640
Db 1085962 GTACCTTTGTTATTGCAAGACGGTATGGCGGAAGATTTAAAGAAAGAGTGGATTAGATT 1086021
Qy 641 CGTACCAATCAATATTGATCTTACAGGAACTATCCAAAGGTGATCTCCCAAGAA 700
Db 1086022 ATTACAAATCAATCTTGATAC-CGCAACAGAAATATTGCAAGAGTATGGTACTGCC 1086080
Qy 701 GCTTTGATGATGATTGAAACATTTCAAAACGTTCAAGGATCTGGATTAAAGGCATGCA 760
Db 1086081 GTTTTGTATGATCGCTTAGCACCTTAGGAAAGTGGCGCAAGCTGGATTAAAGTGTGTT 1086140
Qy 761 CAGTGTGTTATTCTTGGTCTTGGTGAGACCCCAAGAGGACCGGTATCTTTCTTACACT 820
Db 1086141 GTGGTGGCATTGTAGGGATGAATGAACCTCGCAAGAACGGGTGGATTAAATCGGAGCC 1086200
Qy 821 TGGCCCAATGGATCAGATCCAGAGTCTCTTCAATCAACAGACTGTCCCAATCAGG 880
Db 1086201 TTGCAAAATCTTGATCCGCAACCTGAGTCAGTGGCGCGGATTAATCAACTCGTTAAGGTTGAAG 1086260

Qy	881	GCACGCCCAATGTATGAAGAAGTTTAAGAACAGCAAGTCTGAATGTATGATGTGCAGAA	940	LOCATION: (165001)..(180000)	
				OTHER INFORMATION: n=a or c or g or t	
Db	1086261	GTACCGCGTTAGCCGATCGCG-----AAGAATTAGATTGGACGGAGTTTGTGCGGAA	1086311	NAME/KEY: misc feature	
				LOCATION: (180001)..(195000)	
Qy	941	CCATTGCTACTGCAAGATTGGTTCATGCCACGCTCTATTATCAGATTGGCTGCAGGAAGAT	1000	OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
Db	1086312	CGATTGCTGTGTCACGCTATTACTATGCCAAAGATTATGTTCCCTTCTCGAGGTGCGA	1086371	LOCATION: (195001)..(210000)	
				OTHER INFORMATION: n=a or c or g or t	
Qy	1001	ATACAATGAAGAGGCGAGCAACAGGTGATGTGCTTTCATGGCTGGTTGTATATGCCATCTTCA	1060	NAME/KEY: misc feature	
				LOCATION: (210001)..(225000)	
Db	1086372	GTGGTATGACTGAAGAAATGTCAGACGATGTTTTATGCGGGTGCAAATCTTATTTCT	1086431	OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
Qy	1061	CAGTGAAGAAAATGCTCACAACTGTGTAAACGGCTGGGATGAGGATAAGCCATGTTGG	1120	LOCATION: (240001)..(255000)	
				OTHER INFORMATION: n=a or c or g or t	
Db	1086432	ATGCGCATAAATTAATTGTTTACAGATAATCCAGAGAAGATGCGCATCACTAGTTATTGATGG	1086491	NAME/KEY: misc feature	
				LOCATION: (255001)..(270000)	
Qy	1121	CTAAATGGGCTCTGAACCA	1140	OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
Db	1086492	CAAAATTAGATTAGAGCCA	1086511	LOCATION: (270001)..(285000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (285001)..(300000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (300001)..(315000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (315001)..(330000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (330001)..(345000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (345001)..(360000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (360001)..(375000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (375001)..(390000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (390001)..(405000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (405001)..(420000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (420001)..(435000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (435001)..(450000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (450001)..(465000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (465001)..(480000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (480001)..(495000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (495001)..(510000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (510001)..(525000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (525001)..(540000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (540001)..(555000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (555001)..(570000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (570001)..(585000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (585001)..(600000)	
				OTHER INFORMATION: n=a or c or g or t	
				NAME/KEY: misc feature	
				LOCATION: (600001)..(615000)	

```
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (825001)..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
```

```
; NAME/KEY: misc_feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature

Query Match 17.5%; Score 208.2; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 4.5e-54;
Matches 488; Conservative 0; Mismatches 413; Indels 12; Gaps 2;

Qy 206 CCTGGAGCAAGAAGAAATTAAGCTATATATGACACACCACTCATGGAGTTGATGCACT 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208824 CCTGGTCATTAGAAGACATCGCGAAATTTATACACTCCGTTATTTGAGCTGATTCACA 1208765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 ATGCTCAGGTGCAACACAGAGGTTCCAAACACTTCAGAGGTTCAATTTGTGCACTCTTA 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208764 AAGCCAATGCCATATTGCGTAGTAATTTCTCCATTTCAGAACTGCAGACTTGCCTATCTGA 1208705
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 TGAATATCAAACTGGTGGTTGTACCCAGGACTGTAACTACTGTGCCCAATCAGACGTT 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208704 TTTGGAATTAACACTGGTGGATGCGTTGAAGATTTCGCTACTGTGCCCAATCTTTCCCGCT 1208645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 ACAACACTGGTGTCAAGGCTGAAGAATCATCCAAGTTGATGAGTGTATTGAAGTCGCA 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208644 ATCATACCAGCTCACACAGAACCTATGATGAATAATTGTAGAGTTGTGGAAGGGCAA 1208585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 AGGAGCAAGGGCAATGGATCTACAAGGTTCTGTATGGGTGCTGCTTGGAGAGAGATGA 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208584 AACGTGCTGTAGAGTAGGCGCCACTCGTGTGTCTTGGGGCTGCTGGCGCAATGCTA 1208525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 506 AAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCAGTGTCTCATGACCTTG 565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208524 AGG---ACGATCGATACCTTTGATAGAGTCTCGTATGTTGTAAGATATCACAGACTTCG 1208468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 566 GAATGGAGAGTTGTGTACCTCGGAATGTTGATAAAGACCAGCCACTGAAATTTGAAA 625
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208467 GAGCCGAGGTTGTTGTGCTTTTAGGCATGCTCTCCGAGAGCAGCTTAAAAAATGTATG 1208408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 626 GTGCTGGTTGACGGGTACAACCATTAACATTTGATCTTACAAGGAACACTATCCAAAGG 685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208407 ATGCAGGACTTTATGCTCAATCAATTAATTTAGACTCTTCTCGGAATTTCTATGAAACTA 1208348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 686 TGATCTCCACAAGAGCTTTGATGATAGATGAAACATTTCAAAAAGCTTCAAGAGTCTG 745
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208347 TAATCAACAACAGCTTCTTATGAAGATGCGCTCAACACTTTGATGTAGTAATAAATCTG 1208288
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 746 GATTAAAGGCATGCACAGGTGGTATTCTTGGTCTTGGTGAGACCCCAAGAGACCGTCTAT 805
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208287 GCATTAGTACATGCTGGGTGGTATTGTAGGTATGGAGAACTGGAAGAAGACCGTATAA 1208228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 806 CTTTCTCTTACACCTTTGGCCACAATGGATCAGCATCCAGAGTCTCTTTCCAATCAACAGAC 865
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208227 AGCTTCTTCATGTTCTTGTCAACAAGAGATCATATCCAGAAATCCGTACTGTAAATTTAC 1208168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 866 TGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTAAGAACAAGCAAGTTGAAGTTG 925
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208167 TTTGGCCGATTGACGGCAGCCCTTTGCAAGACCAGCCTCCGA-----TTTCTTTCT 1208117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 926 ATGAGATTGTCAACAACATTTGCTACTGCAAGATTGGTCATGCCAACCTCTATTATCAGAT 985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208116 GGGAAAGTCTTCCGAACCATAGCAACGGCAGCGGTGTTTTCCTCCAGATCCATGTAGCAGC 1208057
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 986 TGGCTGCAGGAAGATATACAATGAAAGAGGCAGAACAGGTGTGTGCTTCATGGCTGGTT 1045
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208056 TTGCTGCAGGACGCGCTTTCTCCACAGTAGAACAACAACCTTATGTTTTCTAGCCGGTG 1207997
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1046 GTAATGCCATCTTCACAGGTAAAGAAATGTCTCAACAATGTGTACGGCTGGGATGAGG 1105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1207996 CCAACTCCATATTCTATGGAGATAAACTGTTGACTGTAGAAAAACAATGATATAGATGAAG 1207937
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1106 ATAAAGCCATGTT 1118
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1207936 ATGCTGAAATGAT 1207924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 14

US-09-438-185A-1/c
; Sequence 1, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438.185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1

Query Match 17.5%; Score 208.2; DB 4; Length 1230230;
Best Local Similarity 53.5%; Pred. No. 4.5e-54;
Matches 488; Conservative 0; Mismatches 413; Indels 12; Gaps 2;

QY 206 CTTGGACCAAGAAAGTAAAGCTATATATGACACACCACTCATGACTTGTGACCT 265
DB 1198672 CTTGGTCAATGAAAGCATCCGGAATTTATCACATCCGTTATTTGAGCTGATTCACA 1198613

QY 266 ATGCTCAGGTGCAACAGAGGTTCACAACTTCAGAGGTTCAATTTGTGCACTCTTA 325
DB 1198612 AAGCCCAATGCCATATTGCGTAGTAATTTCTCCATTGAGAACTGCAGACTTGTCTATCTGA 1198553

QY 326 TGAATATCAAACTGTGTGTTTACCGAGCACTGTAAGTACTGTGCCAATCACAGGTT 385
DB 1198552 TTTTCGATTAAGTGTGGATGCGTTGGAAGATTGCGCTACTGTGCCAATCTTCCCGCT 1198493

QY 386 ACAACACTGTGTCAAGCTGAAAGATCATCAAGTTGATGAGTGAATGAAGCTGCA 445
DB 1198492 ATCATCCACGTCACACAGAACCTATGATGAATTTGACGTTGTGGAAGGCA 1198433

QY 446 AGGAGGCAAGGCCAATGGATCTACAAAGGTTCTGTATGGGTGCTGTGGAGAGATGA 505
DB 1198432 AACGTGCTGTAGAGCTAGGCGCCACTCGTGTGTGTCTTGGGGCTGCTGGCGCAATGCTA 1198373

QY 506 AAGGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCATCTGCTTCATGACCTTG 565
DB 1198372 AGG---ACGATCGATATCTTTGATAGAGTCTCCTGATGTTGGAAGATATCACAGATCTCG 1198316

QY 566 GAATGAGAGCTGTGTCAACCTCGGAATGTTGATAAGACCAAGCACTGAATTCGAAA 625
DB 1198315 GAGCCAGGTTTGTGTCTTTAGGCAATGCTCTCCGAGAGCAAGTAAATCTGTATG 1198256

QY 626 GTGCTGGTTGACGGGTACAAACATAATGATATCTTACAAAGGAACACTATCCAAAGG 685
DB 1198255 ATGCAGGACTTTATGCTCAATCAATCAATTAATTTAGACTCTTCTCCGGAATTTATGAAACTA 1198196

QY 686 TGATCTCCCAAGAGCTTTGATGATAGATTGAAACATTCAAAAGCTTCAGGATCTG 745
DB 1198195 TAATCAACACGCTTTTATGAAGATGCGCTCAACACTCTTGTATGATGATTAATAATCTG 1198136

QY 746 GATTAAGGATGACAGAGGTGTTATCTTGGTCTTGTGTAGAGCCCAAGAGGACCGTGTAT 805
DB 1198135 GCATTTAGTACATGCTCGGTGTTTGTAGTATGGAGNATCTGAAGAGACCGTATTA 1198076

QY 806 CTTTCCTCTACACCTTGGCCACAAATGGATCAGCATCCAGAGTCTCTTCCAATCAACAGAC 865

DB 1198075 AGCTTCTTCATGTTCTTGTCAACAAGAGATCATATCCAGAATCCGTACCTGTAAATTTAC 1198016
QY 866 TGGTCCCAATCAAGGCGCAGCCAAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAAGTTG 925
DB 1198015 TTTGGCCGATTGACGCGACGCTTTGCAAGACCAAGCTCCGA-----TTTCTTTCT 1197965
QY 926 ATGAGATTCTCAGAACCAATTGCTTACTGCAAGATTGGTCATGCCAAGCTCTTATTCAGAT 985
DB 1197964 GGGAAAGTCTTGGCAACCATAGCAACGGCACGGGTGTTTTTCCCAAGATCCATGGTACGAC 1197905
QY 986 TGGCTGCAGGAAGATATACAATAAGAGAGGAGAGAGAGTGTGCTTCATGCGCTGGTT 1045
DB 1197904 TTGCTGCAGGACGCGCTTTCCCTCACAGTAGAACAACAACCTTATGTTTCTAGCCGGTG 1197845
QY 1046 GTAATGCCATCTTACAGGTAAAGAAATGCTCACAAATGTGTAACGGCTGGGATGAGG 1105
DB 1197844 CCACTCCATATCTTATGGAGATAAACTGTTGACTGTAGAAAACATGATATAGATGAAG 1197785
QY 1106 ATAAAGCCCATGTT 1118
DB 1197784 ATGCTGAAATGAT 1197772

RESULT 15

US-09-180-109A-23
; Sequence 23, Application US/09180109A
; Patent No. 6410293
; GENERAL INFORMATION:
; APPLICANT: MUKUMOTO, Fujio
; APPLICANT: NISHIO, Shoichi
; APPLICANT: AKIMARU, Jiro
; APPLICANT: MITSUDA, Satoshi
; TITLE OF INVENTION: DNA Fragments Containing Biotin Biosynthetase Gene and
; TITLE OF INVENTION: Use of the Same
; FILE REFERENCE: 0152-0490P
; CURRENT APPLICATION NUMBER: US/09/180,109A
; CURRENT FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 09/047838 JAPAN
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Sphingomonas paucimobilis
; FEATURE:
; OTHER INFORMATION: Strain = JCM7511
; NAME/KEY: CDS
; LOCATION: (223)..(1245)
US-09-180-109A-23

Query Match 17.3%; Score 205.2; DB 3; Length 1420;
Best Local Similarity 52.4%; Pred. No. 5.7e-55;
Matches 509; Conservative 0; Mismatches 448; Indels 15; Gaps 2;

QY 184 TCAGTGAAGACACCACTCAACACCTGGACCAAGAAATTAAGCTATATATGACACA 243
DB 247 TCGAGGCGACCCCGGACCGACTTGAACCCGCGGAGATCGCGCGCTGTTCGACCTG 306

QY 244 CCACTCATGGACTTGATGCACTATGCTCAGGTGCAACAGAGGTTCCAAACACTTCA 303
DB 307 CCTTTCACGAGCTGTGTTCCGCGCGGCGAGGTGACCGCGCATCACCGCGCAT 366

QY 304 GAGGTTCAATTTGTCAGCTCTTATGAATATCAAACTGGTGGTTGTACCGAGCACTGAAG 363
DB 367 CAGGTTCAAGTCTGACGCTGTTGTCGATCAAGACGGGCGGCTGCCCGAGGATTCGCGC 426

QY 364 TACTGTGCCAATCACAGGTTTACAACTGTGTGTCAAGGCTGAAAGAAATCATCAAGTT 423
DB 427 TATTGACGCACTGCAACCCATGCCGATACCGGGCTGAGGCGACCCAGCTGATGACCCG 486

QY 424 GATGAGGTGATTGAAGCTGCAAGAGGCAAGGCAATGGGATCTTCAAGGTTCTGTATG 483

Db 487 CGCGCCGTGTCAGAGGGGGCGCGCAGGCCAAGGATCAGCGCTCGACGCGCTTCTGCATG 546
Qy 484 GGTGCTGCTTGGAGAGAGATGAAGGTAGAAAGTCAAACTTTGAAGAAAATCAAAGAGATG 543
Db 547 GCGCGGCTTGGCGCAACCCCAAGGATCG-----CGACATGCCCGCATCGTGGAGATG 600
Qy 544 ATCACTGCTGTGCATGACCTTGAATGGAGAGTTGTTCACCTGGGAATGGTTGATAAA 603
Db 601 GTGAAGGGGTGCGCGCCATGGCATGGAACCTGATGACGCTGGGCATGCTGACCGAT 660
Qy 604 GACCAAGCCACTGAATTTGAAAGTGTGGTTCACGCGGTACAAACCATACATTTGATACT 663
Db 661 GCACAGCGGAGACGCTCGCCGAGCGGGCTGGACTATTACAAATCAAAATATCGACAG 720
Qy 664 TAAAGGAACACTATCAAAAGGTGATCTCCACAAGAGCTTTTGATAGATAGATTGAATAA 723
Db 721 TCGCCCGAGCGTTATGGCGACGTCATCACACGCGCAGCTTCGGCGAGCGTTGGAGACG 780
Qy 724 TTCAAAACGTTCAAGGATCTGGATTAAAGGCATGCACAGGTGGTATTCTTGGTCTTGGT 783
Db 781 TTGGAGCATGTCCGCGATGCCGCGATCAATGTATGCTGTGGCGGTATTGTGGCATGGGT 840
Qy 784 GAGACCAAGAGACCGTGTATCTTCTCTACACCTTGCCACATGGATCAGCATCCA 843
Db 841 GAGACGCGCGGACCCGGTCCGCTTCATCATGCGCTTGCCACCTGCCGCTCCATCCG 900
Qy 844 GAGTCTCTTCCAATCAACAGACTGTCCCAATCAAGGCGACGCCAATGTATGAAGAAGTT 903
Db 901 GGCAGCGTCCGGTGAACCGCTGGTGCCTGGTCAAGGGCACGGTATTGGGCGATATGTTG 960
Qy 904 AAGAACAAGC-----AAGTTGAAGTTGATGAGATTGTGAGAACCATTCCTACTGCA 954
Db 961 GCCGACACGCCGCTGGCCCAAGATCGACGATATCGAATTCGTCCGACCGCTCGCGTTGCG 1020
Qy 955 AGATTGGTCAATGCCAACGTCATTATATCAGATTGGCTGCAGGAGATATACAAATGAAGAG 1014
Db 1021 CGCATCACCATCCGCGATTTCGATGGTCCGCTGTCCGCGGGGCGGAGAGCATGTCGGAT 1080
Qy 1015 GCAGAACAGGTGATGTGCTTCATGGTGTGTTGTAATGCCATCTTACAGGTAAGAAAATG 1074
Db 1081 GCCACCCAGGCTTTGTGCTTCTGGCGGCGGAACTCGATCTTACCGGCGACAAGCTG 1140
Qy 1075 CTCACAACAAATGTAAACGCTGGGATGAGGATAAAGCCATCTTGGCTAAATGGGGTCTG 1134
Db 1141 CTGACTGGGGCAATGCGGGCGACGACAAAGGACGAGCGCTCTTCGCCGCGCTGGGGCTC 1200
Qy 1135 AAACCAATGGAG 1146
Db 1201 ACGCCCATGGCG 1212

Search completed: September 17, 2005, 20:39:10
Job time : 251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 17:54:15 ; Search time 912 Seconds
(without alignments)
8678.876 Million cell updates/sec

Title: US-09-752-957F-1

Perfect score: 1188

Sequence: 1 atgctgtttattgactgc.....aggttgattcggtgcttga 1188

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	21	US-09-752-957B-1
2	417.6	35.2	1125	21	US-10-741-849-6038
3	386.2	32.5	1334	18	US-10-425-114-28355
4	386.2	32.5	1345	20	US-10-425-115-180438
5	349.2	29.4	1659	9	US-09-740-288A-29
6	349.2	29.4	1659	16	US-10-431-544-29
7	328.2	27.6	3083	18	US-10-390-472-1
					Sequence 1, Appli
					Sequence 6038, Ap
					Sequence 28355, A
					Sequence 180438,
					Sequence 29, Appli
					Sequence 29, Appli
					Sequence 1, Appli

8	318.2	26.8	1420	19	US-10-437-963-18311	Sequence 18311, A
9	311.8	26.2	1515	9	US-09-740-288A-21	Sequence 21, Appli
10	311.8	26.2	1515	16	US-10-431-544-21	Sequence 21, Appli
11	310.2	26.1	1396	9	US-09-740-288A-17	Sequence 17, Appli
12	310.2	26.1	1396	16	US-10-431-544-17	Sequence 17, Appli
13	310.2	26.1	1439	9	US-09-740-288A-23	Sequence 23, Appli
14	310.2	26.1	1439	16	US-10-431-544-23	Sequence 23, Appli
15	310.2	26.1	1486	18	US-10-425-114-28968	Sequence 28968, A
16	310.2	26.1	1514	20	US-10-425-115-71987	Sequence 71987, A
17	306.6	25.8	1152	9	US-09-740-288A-7	Sequence 7, Appli
18	306.6	25.8	1152	16	US-10-431-544-7	Sequence 7, Appli
19	305.4	25.7	1443	19	US-10-767-701-13006	Sequence 13006, A
20	304	25.6	1137	9	US-09-938-842A-579	Sequence 579, App
21	304	25.6	1137	11	US-09-938-842A-579	Sequence 579, App
22	304	25.6	1137	21	US-10-702-364-1	Sequence 1, Appli
23	295.6	24.9	1477	9	US-09-740-288A-25	Sequence 25, Appli
24	295.6	24.9	1477	16	US-10-431-544-25	Sequence 25, Appli
25	294.4	24.8	1340	9	US-09-740-288A-11	Sequence 11, Appli
26	294.4	24.8	1340	16	US-10-431-544-11	Sequence 27, Appli
27	292.8	24.6	1526	9	US-09-740-288A-27	Sequence 27, Appli
28	292.8	24.6	1526	16	US-10-431-544-27	Sequence 27, Appli
29	292.2	24.6	1370	18	US-10-425-114-7490	Sequence 7490, Ap
30	292.2	24.6	1614	18	US-10-424-599-24740	Sequence 24740, A
31	279.8	23.6	1111	18	US-10-425-114-10388	Sequence 10388, A
32	260.8	22.0	1467	9	US-09-740-288A-19	Sequence 19, Appli
33	260.8	22.0	1467	16	US-10-431-544-19	Sequence 19, Appli
34	243.4	20.5	1032	9	US-09-740-288A-31	Sequence 31, Appli
35	243.4	20.5	1032	16	US-10-431-544-31	Sequence 31, Appli
36	237.6	20.0	987	17	US-10-282-122A-8411	Sequence 8411, Ap
37	228.6	19.2	957	17	US-10-282-122A-25092	Sequence 25092, A
38	223.6	18.8	1002	9	US-09-815-342-7073	Sequence 7073, Ap
39	223.6	18.8	1002	17	US-10-282-122A-22203	Sequence 22203, A
40	212.6	17.9	1038	17	US-10-282-122A-42017	Sequence 42017, A
41	211.6	17.8	1830121	17	US-10-329-670-1	Sequence 1, Appli
42	211.6	17.8	1830121	20	US-10-158-865-1	Sequence 1, Appli
43	211.6	17.8	1830121	22	US-10-981-687-1	Sequence 1, Appli
44	210.4	17.7	1907	18	US-10-425-114-14512	Sequence 14512, A
45	208.2	17.5	996	17	US-10-312-273-284	Sequence 284, App

ALIGNMENTS

RESULT 1

US-09-752-957B-1
; Sequence 1, Application US/09752957B
; Publication No. US20030104584A1
; GENERAL INFORMATION:
; APPLICANT: David SHUAN
; TITLE OF INVENTION: YEAST WITH HIGH BIOTIN PRODUCTIVITY AND THE PREPARATION METHOD 1
; FILE REFERENCE: C. utilis BIO2 - biotin synthase
; CURRENT APPLICATION NUMBER: US/09/752,957B
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Candida utilis
US-09-752-957B-1

Query Match 100.0%; Score 1188; DB 10; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGTCGTTTATTGACTGCTATTAGTCGTCGAGTCTCTTCCACTCTTAGAGTAGCT 60
Db	1	ATGTCGTTTATTGACTGCTATTAGTCGTCGAGTCTCTTCCACTCTTAGAGTAGCT 60
Qy	61	TCTAGGCTACTCTGGCAACAGGTGCTGCTGCTCGGAGATCTTGGAGATGTGTTTC 120
Db	61	TCTAGGCTACTCTGGCAACAGGTGCTGCTGCTGCTCGGAGATCTTGGAGATGTGTTTC 120

```
121 ACAGAACTTGAAGAGTGGCTTTCACAGGAGAGAGCCAAACCCATTGGAATATGCA 180
121 ACAGAACTTGAAGAGTGGCTTTCACAGGAGAGAGCCAAACCCATTGGAATATGCA 180
181 TTGTCACTGAAGACACAGTCAACACCTGACACCAAGAGAAATTAAGCTATATATGAC 240
181 TTGTCACTGAAGACACAGTCAACACCTGACACCAAGAGAAATTAAGCTATATATGAC 240
241 ACACCACTCATGACTTGTATGCTATGCTCAGTGCACACACAGAGGTTCCAAACACCT 300
241 ACACCACTCATGACTTGTATGCTATGCTCAGTGCACACACAGAGGTTCCAAACACCT 300
301 TCAGAGGTTCAATGTGCACTTATGAATATCAAACTGGTGGTTGTACCGAGGACTGT 360
301 TCAGAGGTTCAATGTGCACTTATGAATATCAAACTGGTGGTTGTACCGAGGACTGT 360
361 AAGTACTGTGCCAATCACAGCGTTACAACTGGTGTCAAGCTGAAGAAATCATCCAA 420
361 AAGTACTGTGCCAATCACAGCGTTACAACTGGTGTCAAGCTGAAGAAATCATCCAA 420
421 GTTGATGAGGTGATTGAAGCTGCAAGGAGGCAAGGCCAATGGATCTACAAGGTTCTGT 480
421 GTTGATGAGGTGATTGAAGCTGCAAGGAGGCAAGGCCAATGGATCTACAAGGTTCTGT 480
481 ATGGGTGCTGCTTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAAGAG 540
481 ATGGGTGCTGCTTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAAGAG 540
541 ATGATCACTGCTGTCATGACCTTGGAAATGGAGAGTTGTGTACCTCGGGAATGGTTGAT 600
541 ATGATCACTGCTGTCATGACCTTGGAAATGGAGAGTTGTGTACCTCGGGAATGGTTGAT 600
601 AAGACCAACGCCACTGAATTTGAAGGCTGGGTTGACGGCTACACCAATTAACATTTGAT 660
601 AAGACCAACGCCACTGAATTTGAAGGCTGGGTTGACGGCTACACCAATTAACATTTGAT 660
661 ACTTACAAGGAACACTATCCAAAGGTGATCTCCACAAAGAGCTTTGTATGATGATGAA 720
661 ACTTACAAGGAACACTATCCAAAGGTGATCTCCACAAAGAGCTTTGTATGATGATGAA 720
721 ACATTCAAAAAGCTTCAAGATCTGGATTAAAGGCATGCAAGGTGGTATCTTGTGCTTT 780
721 ACATTCAAAAAGCTTCAAGATCTGGATTAAAGGCATGCAAGGTGGTATCTTGTGCTTT 780
781 GGTGAGACCCAAAGGACCGTGTATCTTCTCTACACCTTGGCCACAAATGGATCAGCAT 840
781 GGTGAGACCCAAAGGACCGTGTATCTTCTCTACACCTTGGCCACAAATGGATCAGCAT 840
841 CCAGAGTCTCTTCCAATCAACAGACTGGTCCCAATCAAGGGCAGCGCCCAATGATGAAGAA 900
841 CCAGAGTCTCTTCCAATCAACAGACTGGTCCCAATCAAGGGCAGCGCCCAATGATGAAGAA 900
901 GTTAAGAACCAAGCTGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 960
901 GTTAAGAACCAAGCTGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 960
961 GTCAATGCCAAGCTCTATTATCAGATTGGCTGCAGGAGATATACAAATGAAGAGGAGAA 1020
961 GTCAATGCCAAGCTCTATTATCAGATTGGCTGCAGGAGATATACAAATGAAGAGGAGAA 1020
1021 CAGGTGATGCTCTTCAATCAACAGACTGGTCCCAATCAAGGGCAGCGCCCAATGATGA 1080
1021 CAGGTGATGCTCTTCAATCAACAGACTGGTCCCAATCAAGGGCAGCGCCCAATGATGA 1080
1081 ACAATGTGTAACCGCTGGGATGAGGATTAAGCCATGTTGGCTAAATGGGGTCTGAACCA 1140
1081 ACAATGTGTAACCGCTGGGATGAGGATTAAGCCATGTTGGCTAAATGGGGTCTGAACCA 1140
1141 ATGGAGAGTTTCAAAATCAAAACCAAGGAGGTTGATGCTTGGTGGCTTGA 1188
1141 ATGGAGAGTTTCAAAATCAAAACCAAGGAGGTTGATGCTTGGTGGCTTGA 1188
```

```
RESULT 2
US-10-741-849-6038
; Sequence 6038, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 10182-023-999
; CURRENT APPLICATION NUMBER: US/10/741,849
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/434,832
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6038
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-741-849-6038
```

Query Match 35.2%; Score 417.6; DB 21; Length 1125;
Best Local Similarity 67.0%; Pred. No. 8.4e-116;
Matches 645; Conservative 0; Mismatches 299; Indels 18; Gaps 3;

```
QY 207 CTGGACCAAGAGAAATTAAGCTATATACACACACCACTCATGGACTTGTGCACTA 266
DB 159 CTGGACCAAGAGAAATTAAGCTATATACACACACCACTCATGGACTTGTGCACTA 218
QY 267 TGTCTCAGGTGCAACACAGAGGTTCCAAACCTTTCAGAGGTTCAATTTGTGCACTTAT 326
DB 219 CGTCAATTTGAACATCTGTGAGTACCAAAAACCTGCTGAAGTCCCAATTTGTGCT 278
QY 327 GAATATCAAAACCTGGTGTGTACCGAGGACTGTGAAGTACTGTGCGCAATCAAGGCTTA 386
DB 279 TAACATTAAGCAGGGTGGATGTACCGAAAACCTGTTTACTGTTCTCAATCATCGAACA 338
QY 387 CAACACTGCTGTCAGGCTGAAGATCATCAAGTTGATGAGGTTGAAGCTCAAA 446
DB 339 TGATCTGAGGTTCAAGCTGAAGAAACCTGCTGATGCTGATGAGCAATTTGCTAG 395
QY 447 CGAGGCAAAAGGCAATGGAATCTCAAGGTTCTGTATGGTGTGCTTGGAGAGATGAA 506
DB 396 AGATGCAAGGACAGGAGTGTGCTAGAGTTTGGCTTGGCTTGGAGAGATGCA 455
QY 507 AGGTGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTGTCATGAC---CT 563
DB 456 AGGAAGAAAGTCTGCCCTTAGACAAAATCAGTAAAGTTGTTAAACAAATCAACGACGAT 515
QY 564 TGGATGAGAGCTGTGTCACCTCGGAATGTTGATTAAGACCAAGCCACTGAATTGAA 623
DB 516 GAAGTTGGAAGTCTGTGTTACTCTCGGATGATCAACGAAAGAGGCTGAAATTTTAA 575
QY 624 AAGTGTGCTGCTGAGGCTGACCAACATTAACATTTGATCTTACAGGAAACACTATCCAA 683
DB 576 ACAGATGCTGTTGCTGCTTACCAACCAACATTTGACACTTCAAGAGAACATTTATCAA 635
QY 684 GGTGATCTCCAAAGAGGCTTTGATGATAGATTGAATAACATTTCAAAAAGCTTCAAGGATC 743
DB 636 CGTGTGTCACCAACCGGTACATAGACGAAAGATTAGAAACCATTTAAACAGCTCCAAAAGC 695
QY 744 TGGATTAAGGCTGACACAGGTTGTTCTTGGTCTTGGTGGAGCCCAAGAGGACGCTGT 803
DB 696 TGGTATCAAGGCTTGTACCGGTGGTATTCTCGGATTTGGGAGAAACCGCAAGATCATGT 755
QY 804 ATCTTTCTCTACACCTTGGCCCAATGATGATCAGCATCCAGAGTCTCTTCCCAATCAACAG 863
DB 756 TTGCTTTTGTACACTTTGTCCCAATGATGTCGACCCAGAAATCATTTGCCAATCAACAG 815
QY 864 ACTGTGTCCTCAATCAAGGGCAGCGCAATGATGAA-----GAAGTTAAGAACAA 911
```

Db 816 ATTAGTTCCAAATCAGGGTACTCCAAATTTGTTGAAGTGTGAAAAATGCACCAAGGAAAG 875
Qy 912 GCAAGTTGAAGTTGATGAGATTGTGAGAACATTGCTACTGCAAGATTGGTCAATGCCAAC 971
Db 876 ACAATTGGAATTTGATGCTATTTTGAGAACATTTGCTACTGCTAGATTGATTATGCTTGA 935
Qy 972 GTCTATTATCAGATTGGTGTGAGGAAGATATACAAATGAAAGAGGAGAGACAGGTGATGTG 1031
Db 936 ATCCATCATTAAGATTGGCGCGGTGATACACCATGAAGGAACACGAGCAGTTCTTTGTG 995
Qy 1032 CTTCAATGGCTGTTGTTAATGCCATCTTCAAGGTAAGAAATGCTCAACAATGTGTAA 1091
Db 996 TTTTCATGAGTGGGTCAATGCTATTTTACCGGTAAAGAGATGTTGACTACAATGTGTAA 1055
Qy 1092 CGGCTGGGATCAGCATAAAGCCATGTTGGCTAAATGGGTCTGAAACCAATGGAGAGTTT 1151
Db 1056 TGGTTGGGATGAAGATATCGCCATGTTGAAGAAATGGGTTTGAACCAATGGAAAGTTT 1115
Qy 1152 CA 1153
Db 1116 CA 1117

RESULT 3

US-10-425-114-28355
; Sequence 28355, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28355
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4756-066-B10_FLI
US-10-425-114-28355

Query Match 32.5%; Score 386.2; DB 18; Length 1334;
Best Local Similarity 63.0%; Pred. No. 3.4e-106;
Matches 617; Conservative 0; Mismatches 353; Indels 9; Gaps 1;
Qy 169 TTGGAATGATGATGTTGAGTGAAGACACCACTGACCTGAGCAAGAAAGAAATTTAA 228
Db 170 TTTGAGAATGCACTCAATGCAAGACACCGCGGCAAAATGGAAGAGGAGATTTCA 229
Qy 229 GCTATATATGACACCACTCATGACTTGTGATGACTATGCTCAGGTGCAACAGAGG 288
Db 230 GAGATATACAACTCTCGCTGATTGACCTTACCTATGCTTCGGCTTGGTCCACCGCGA 289
Qy 289 TTCCAAACAACTTCAGAGTTCAATTTGCTACTCTTATGAATATCAAACTGGTGGTTGT 348
Db 290 TTTCCATGACCCGCTGCCATTCAGATGTATACCTCTTCAACATCAAGCGGGCGGTGC 349
Qy 349 ACCGAGGACTGTAAAGTACTGTGCCCAATTCAGAGGTTTACAACTGGTGTCAAGGCTGAA 408
Db 350 AGTGAAGACTGTTCTTACTGCGCCAGTCATCTCGTAGGACACCGGTCTCAAGGCACA 409
Qy 409 AGAATCATCCAAAGTTGATGAGTGAATGAACTGCAAGAGGCAAGGCAATGGATCT 468
Db 410 AAGCTCAGCTCTGTGCAATCGTCTCGAAGCTGCAAGCATAGCTGCAAGAAACCGGAGC 469

Qy 469 ACAAGGTTCTGTATGGTCTGCTTGGAGAGAGATGAAGGTAGAAAGTCAAAGTCAAAGTCAAG 528
Db 470 TCAGATTTCTGCAATGGCGCGCGTGGAGAGACATGAGGGGTCTGTAAGACGAACCTCAAG 529
Qy 529 AAAATCAAAGAGATGATCACTGCTGTCCATGACCTTGGAAATGAGAGATTGTGTCAACCTG 588
Db 530 AACATCAAAGAAATGATCAAGGGGTGCGGGCATGGGCATGGAGGATCGGTGACCTT 589
Qy 589 GGAATGGTTGATAAAGACCAAGCCACTGAATGAAAGTCTGCTGGTTGAGCGGTGACCAAC 648
Db 590 GGCATGCTGATGCTCAAGCAAGAACTGAAGATCGGGCTTGAAGGGCTTACAAAT 649
Qy 649 CATAACATGATGATCTTACAAGGAACACTATCCAAAGGTGATCTCCCAAGAAAGCTTTGAT 708
Db 650 CACAATGTGATGATGAGCGCGGAGCACTACCTAGCGTCAATCCACACAGCTAGTACGAC 709
Qy 709 GATAGATTGAAAAACATTCAAAAAAGTTCAAGGATCTGGAATTAAGAGGATGACACAGGTGT 768
Db 710 GAGGACTCAACACCATCAAGAAATGTCAGGAAGCAGGTATTACGTTGTGACCGCGGT 769
Qy 769 ATTCTTTGGTCTTGGTGAGACCCCAAGAGGACCGTGTATCTTTCTTACACCTTGGCCACA 828
Db 770 ATTCTTTGGTCTGCGGGAAGAAAGCCCGTGACCAAGTGGTCTCATCCACACTGTAGCTACA 829
Qy 829 ATGGATCAGCATCCAGAGTCTTTCNAATCAACAGACTGTGCTCCCAATCAAGGGCACGCCA 888
Db 830 CTACCGCGCCATCCAGAGTCTTTCCCGCTCAAGCCCTTGTCCCATCAAGAGGCAACCCCA 889
Qy 889 ATGTATGAAGAAGTTAAGAACAAAGAGTTGAAGTTGATGATGATGTCAGAACCATTTGCT 948
Db 890 CT-----AGGAGAAACCCCAAGCAATCTCTTTGATGCCATCTTCAGAACCATTTGCT 940
Qy 949 ACTGCAAGATTGGTTCATGCCAAGCTCTATTATATCAGATTGGCTGCAGGAAGATATCAATG 1008
Db 941 ACTGCAAGCTCTGCTGCTTCCACCAACCATCATCGTCTCGCAGCGCGCCACCAATG 1000
Qy 1009 AAAGAGCAGAACAGGTGATGCTGCTTCAATGCTGGTGTGTTGATGCTATTCACAGGTAG 1068
Db 1001 CGTGAGGAAAAACAGATTCTGTGCTTCCAGGCTGGTGCAACCGCGTTCCTCTGCGGAG 1060
Qy 1069 AAAATGCTCAACAATGTTGAAGCTGGGATGAGATGAAGATGAAGATGTTGGCTAAATGG 1128
Db 1061 AAGATGCTCAACACCGCGTGTATGTTGGTGAAGAAAGCAAGGCTATGTTTCAGAGGTGG 1120
Qy 1129 GGTCTGAAACCAATGGAGA 1147
Db 1121 GGGCTTAGGCCAATGAAGA 1139

RESULT 4

US-10-425-115-180438
; Sequence 180438, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180438
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9613C.1
US-10-425-115-180438

Query Match 32.5%; Score 386.2; DB 20; Length 1345;

Best Local Similarity 63.0%; Pred. No. 3.4e-106;
Matches 617; Conservative 0; Mismatches 353; Indels 9; Gaps 1;

169 TTGGAATATGATGTCAGTGAAGACCACTCAACCTCGACCAAGAGAAATATAA 228
181 TTTGAGATGACATCAATGCAAGGACCGCGGCAAAATGGCAAGAGGAGATTCA 240
229 GCTATATATGACACACCACTCATGACTTGTATGACTATGCTCGAGTGCACACAGAGG 288
241 GAGATATACAACATTCGCTGATGACCTTACCTATGCTTGGCCCTTGGTCCACGCGCA 300
289 TTCCAAACACCTTCAGAGGTTCAATGTGCACTCTTATGAATATCAAACTGGTGGTGT 348
301 TTCCATGACCCGCTGCCATTCAGATGTATGCTTCAACATCAAGACGGCGGTGTC 360
349 ACCGAGACTGTAGTACTGTGCCAATCAAGGTTTACACACTGGTGTCAAGGCTGAA 408
361 AGTGAAGACTTCTCTACTGCGCCAGTCTCTCGCTACGACCGGTCTCAAAAGCAC 420
409 AGAATCATCCAAGTTGATGAGTGAATGAAGCTGCAAGGAGGCAAGGCCAATGGATCT 468
421 AAGCTCAGCTCTGCGATCCGTCCTCGAAGCTGCACCGCATAGCTAAAGAAAACGGCAGC 480
469 ACAAGGTTCTGTATGGGTGCTGCTGGAGAGAGATGAAGGTAGAAATCAAACTTGAAG 528
481 TCGAGATTCTGATGGCGCGCTGGAGAGACATGAGGGTCTGTAGACGAACTCAAG 540
529 AAAATCAAGAGATGATCAGTCTGTCCATGACCTTGGAAATGGAGAGTTGTGTCACCTG 588
541 AACATCAAGAAATGATCAAGGGGTGCGCGCATGGGCATGGGATGCGTGACCCCTT 600
589 GGAATGTTGATAAGACCAAGCACTGAATGAAAGTGTGCTGACGCGGTGACCAAC 648
601 GGCATGTCGATGCTCAAGCAAGAAAGTGAAGATGCGGGCTTGACGGCTTACAA 660
649 CATAACATTTGATCTTCAAGGAACACTATCCAAAGGTGATCTCCACAAAGCTTTGAT 708
661 CACAATGCTGATACGCGCGGAGCACTACCTTAGCGTCAACACACAGTACGTACGAC 720
709 GATAGATTGAAAACATTTCAAAAAGTTCAAGGATCTGGAATTAAGGCAATGACAGTGGT 768
721 GAGCGACTCAACACCATCAAGAATGTCAGGAAGCAGGTATTCAGTGTGACCGGGGT 780
769 ATTCTTGGTCTGTGAGACCAAGAGACCGGTGATCTTCTTCTACCTTACCTTGGCCACA 828
781 ATTCTTGGTCTGTGAGGAAAAAGCCGTCACCGTGGTCTCTATCCACACTGTAGTACA 840
829 ATGGATCAGATCCAGAGTCTCTTCCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCA 888
841 CTACCGGCCATCAGAGTCTCTTCCCGTCAAGCCCTTGTCCCATCAAGGCAACCCCA 900
889 ATGTATGAAGAGTTAAGAACAGCAAGTGTGAAGTGTGATGATGTTGAGAACCATTTGCT 948
901 CT-----AGGAGAAACCCAAAGCCATCTCTTGTATGTCATCTCAGAACCATTTGCT 951
949 ACTGCAAGTTGGTCATGCCAAGCTTATATCAGATTGGTGGTGCAGGAATATACATG 1008
952 ACTGCAAGCTCGTCTTCCACCAACCATCATCGTCTCGCAGCGCGGCCCAACCATG 1011
1009 AAAGAGCAGAACAGGTGATGCTTCAATGCTGGTGTGATGTCATCTTCAAGGTAAAG 1068
1012 CGTGAAGAAAACAGATTCGTGCTTCAGGCTGTGCAACCGCGTTCACCTGGCGAG 1071
1069 AAAATGCTCAACAATGTGTAAACGGGTGGATGAGGATTAAGCCATGTTGGCTAAATGG 1128
1072 AAGATGCTCAACACGGGTGTAATGGGTGGGAAAGAACAGAGGCTATGTTTTCAGAGGTG 1131
1129 GGTCTGAACCAATGGAGA 1147
1132 GGGCTTAGGCCAATGAAGA 1150

RESULT 5

US-09-740-288A-29
; Sequence 29, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Glycine max
US-09-740-288A-29

Query Match 29.4%; Score 349.2; DB 9; Length 1659;
Best Local Similarity 60.6%; Pred. No. 7.1e-95;
Matches 595; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

173 AATATGCAATGTGTCAGTGAAGACACCACTCAACCTGGACCAAGAGAAATTAAGCTA 232
234 AAGAAGCACTCAACGCAACTGAACCCCGCAGCAATTTGGACAGAGAAATTAAGCGGA 293
233 TATATGACACACCACTCATGCACTTGTATGCACTATGCTCAGGTGCAACACAGAGGTTCC 292
294 TCTATGATAGCCATTGATGAGTATTTGGGGTGTGCTAGTTTGCACAGGAAATTC 353
293 AACAACTTCAGAGGTTCAATTTGTGCACTCTTATGAATATCAAACTGGTGGTTGTACCG 352
354 ATATACCTGGGGCTATTTCAGATGTGTACATTTGTGAACATCAAGACGGGTGGTTCGG 413
353 AGGACTGTAGTACTGTGCCCCAATCAAGGTTTACACACTGGTGTCAAGCTGAAAGAA 412
414 AGGATTTCTTACTGGCCCAATCATCCGCTACCAACCGGTCTCAAGAGCTTCCAAA 473
413 TCATCCAAAGTTGATGAGGTGATTTGAAGCTGCAAGAGGCAAGGCCAATGGATCTCAA 472
474 TGGTCTCCGTCGATCTGCTCGACGCCGCCGATCGCCAAAGACACAGGTAGTACAC 533
473 GGTTCGTATGGTGTGCTGTTGGAGAGAGATGAAGAGTGAAGTCAAACCTTGAGAAAA 532
534 GTTCTGATGGGAGCGCGTGGCGGATATGCGTGGACGAAAAACCAATCTCAAAAATG 593
533 TCAAGAGATGATCACTGCTGCCATGACCTTGGAAATGGAGAGTTGTGTCACCTCGGAA 592
594 TCAAAACAATGGTTAGCGAGATTTCGGGAAATGGGTATGGAAGTATGTGTCAGCTTGGTA 653
593 TGGTGTATAAGAACCAAGCCACTGAAATGAAAAAGTCTGGGTTGACGGCGTACAAACATA 652
654 TGATTTGTCAGAGCAAGCTCAGGAATCAAGAGCCGCTCTCACGGCTTATAATCAT 713
653 ACATTTGATCTTCAAGGAACACTATCCAAAGGTGATCTCCAAAGAAAGCTTTGATGATA 712
714 ATGTGGATACGTCGAGGATTTCTATCCCAAGGTTATCAGACACGAGCACTTATGATGAGA 773
713 GATTGAAACATTTCAAAAGGTTTCAGGATCTGGATTAAGGCATGACAGAGGTGATTC 772
774 GATTGGATACCATTAAGAATGTGAGAGACCGGAAATCAATGTTGTACGGGTGGAATTC 833
773 TTGCTCTTGTGTGAGACCAAGAGACCGTGTATCTTTCTCTACCTCTGCGCCACAATGG 832
834 TCGGATTAGAGAGAAATAAGCTTGACCATATTGGACTTTTGGAGACGGTTGCTACGTTGC 893
833 ATCAGCATCCAGAGTCTCTTCCCAATCAACAGATGTTGTCCTCAATCAAGGGCAACGCAATGT 892
894 CTTGCGATCCGGAATCAATTTCTCTGTGAACATGTTAGTGGCTATCAAGAGAAACCACTGG 953

Db 685 GCTGGAATAACATATGCTCAGGTGGAATTAATTGGTCTTGGAGAGGCGAGAGGACCGT 744
Qy 802 GTATCTTTCTCTACACCTTGGCCACCAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
Db 745 GTAGGATGTTGCATACACTGGTACTTTGGCTGCACACCCAGAGAGTGTTCCTATTAAAT 804
Qy 862 AGACTGGTCCCAATCAAGGGCAGCGCAATGATGATGAAGAGTTAAGAACCAAGTTGAA 921
Db 805 GCATTGGTGTGTCAAGAGGCGCCCTTCAAG-----ATCAGAGCCCTGTAGAG 855
Qy 922 GTTCATGAGATTGTCAGAACCATCTGCTACTGCAAGATTGGTCATGCCAACTCTATTATC 981
Db 856 ATCTGGGAAATGATCCCGCATGATTCGCCACAGCTCGAATGTGATGCCAGAGCAATGGTG 915
Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGCGAGAACAGGAGTGTGCTTCATGGCT 1041
Db 916 AGGCTTTCAGCAGACGAGTAGCTTTCTCCATGCCAGAGAGCGCTCTGCTTCTTAGCT 975
Qy 1042 GGTGTAAATGCATCTTTCAGAGTAAGAAATGCTCACAACAATGTGTAAACGGCTGGGAT 1101
Db 976 GGGGCCAACTCCATCTTCGCTGGTGAGAAACTTCTTAAACACCAAAACACAGCACTTTGAT 1035
Qy 1102 GAGATTAAGCCATGTTGGCTTAATGGGCTCTGAAACCAATGG 1144
Db 1036 GCAGACCAGGCGATGTTCAAGATCCTTGGTCTGATTCCTCAAGG 1078

RESULT 9

US-09-740-288A-21
; Sequence 21, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; TYPE: DNA
; ORGANISM: Zea mays
US-09-740-288A-21

Query Match 26.2%; Score 311.8; DB 9; Length 1515;
Best Local Similarity 59.1%; Pred. No. 1.6e-83;
Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 251 AACCACTGGAGCCGGCCGAGATCCAGCCGCTTACGACTACCGCTCTCGACTCTC 310
Qy 262 CACTATGCTAGGTGCAACACAGAGGTTCCAAACCTTCAGAGGTTCAATGTGCACT 321
Db 311 TTTCACGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAGTGCAGCAATGCACA 370
Qy 322 CTTATGATATCAAACTGGTGTGTACCGAGGACTGTAGTACTGTGCCAATCACAG 381
Db 371 CTTCTTTTCAATCAAGACTGGTGGATGTCAGTGAAGATTGTTCTTACTGTCTCAGTCATCA 430
Qy 382 CGTTACAACACTGGTGTCAAGCTGAAAGAAATCATCCAAGTTGATGAGGTGATTGAAGCT 441
Db 431 AGATACAACTGATGATGAAGGCCCAAAATTTGATGAAGAATGCTGTCTTGAAGCA 490
Qy 442 GCAAGGAGGCAAGGCCAATGGATCTTACAAGGTTCTGTATGGGTGCTGTGGAGAGAG 501
Db 491 GCAAAAAGGCAAAAGAGTCTGGGAGCACCGCTTTTTCATGGGAGCTGTCATGGAGAGAA 550

Qy 502 ATGAAAGGTAGAAAGTCAAACTTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGAC 561
Db 551 ACCATTGGCAGGAATCAAACTTTCAACAGATTTCTTGAATATGTCAAGGAAATTAAGGGGT 610
Qy 562 CTTGGAATGAGAGTGTGTGCACCTGGGAATGTTGATAAAGACCAAGCCACTGAAATTG 621
Db 611 ATGGGATGAGAGTGTGTGTGCACACTAGGATGATAGAAACCAACAGCTGAAGAACTC 670
Qy 622 AAAAGTCTGGTGTGACGGGTACAACTAACATTTGATTAATCAAGGAAACATATCCA 681
Db 671 AAGAAGCTGGACTTACAGCATATAATCAATCAACCTAGATACATCAAGAGAGTATTATCCC 730
Qy 682 AAGTGAATCTCCACAGAGCTTTGATGATAGATTGAAACATTTCAAAAAGCTTCAAGGA 741
Db 731 AACATTATTACCAAGATCATATGATGATAGACTGTCAGACTCTTTGAGCATGTCCGTGAA 790
Qy 742 TCTGGATTAAGGCATGCACAGGTGGTATTCTTGGTCTTGGTGAGACCCAAAGAGACCGT 801
Db 791 GCTGGAATAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAGAGAGAGACCGG 850
Qy 802 GTATCTTTCTCTACACCTTTGGGCACAAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
Db 851 GTAGGTTGTTGCATACCTTAGCTACCTTGCCTACACACCCAGAGAGGTTCTCTATTAA 910
Qy 862 AGACTGTGTCCTCAATCAAGGGCAGCCCAATGTATGAAGAGTTAAGAACCAAGTTGAA 921
Db 911 GCATTGGTGTGTAAAGGCACACCTCTTTGAG-----ACCAGAGCCTGTAGAG 961
Qy 922 GTTGTAGAGATTGTCAGAACCATTTGCTACTGCAAGATTGGTCATGCCCACTGATTATC 981
Db 962 ATCTGGGAATGATCCGCATGTCGCCACTGCTCGGATCAGATGCCAAAGGCAATGGTG 1021
Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGCGACAGACAGGTGATGTGCTTCATGGCT 1041
Db 1022 AGCTTTTCAGAGCGCCGAGTACGGTCTTCGATGCCAGAACAGCGCTGTGCTTCTCGCT 1081
Qy 1042 GGTGTAAATGCCATCTTTCACAGGTAAAGAAATGCTCAACAATGTGTAAACGGCTGGAT 1101
Db 1082 GGGGCCAACTCCATCTTTGGCGCGAGAAACTTCTCAACACCCGCAACACGACTTTGAT 1141
Qy 1102 GAGGATTAAGCCATGTTGGCTTAATGGGCTCTGAAACCAATGG 1144
Db 1142 GCGAGCCAGGCGATGTTCAAGATCCTTGGCTGATCCCAAGG 1184

RESULT 10

US-10-431-544-21
; Sequence 21, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Zea mays
US-10-431-544-21

Query Match 26.2%; Score 311.8; DB 16; Length 1515;
Best Local Similarity 59.1%; Pred. No. 1.6e-83;
Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;

202 AACACCTGGACCAAGAGAAATTAAGCTATATATATGACACACACCTCATGGACTTGATG 261
251 AACGCTGGAGCGCGCGGAGATCCAGGCGTCTACGACTCACGCTCTCGACCTCTC 310
262 CACTATGCTCAGGTGCAACACAGAGGTTTCAACACCTTCAGAGGTTCAATTTGTCAC 321
311 TTTCACGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAGTGCAGCAATGCACA 370
322 CTTATGAATACAACTGGTGTGTACGAGGACTGTAAGTACTGTGCCCCAATCACAG 381
371 CTTCTTTCAATCAAGACTGGTGTGATGAGTGTGTTTACTGTCTCAGTCATCA 430
382 CGTTACAACACTGGTGTCAAGGCTGAAGAAATCATCCAAGTGTATGAGGTTGAAGCT 441
431 AGATACAACACTGGATGAAGGCCCAAAATTTGATGAACAAAGATGTGTCTTGGAGCA 490
442 GCAAGGAGCAAGGCAATGATCTACAAGTTCTGTATGGTGTGCTTGGAGAGAG 501
491 GCAAAAAGGCAAAAGAGTCTGGAGCACCGGTTTTTGCATGGAGGCTGCATGGAGAGAA 550
502 ATGAAAGGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTCCATGAC 561
551 ACCATGGCAGGAATCAAACTTCAACAGATCTTGAATATGTCAAGGAATAAGGGGT 610
562 CTTGGATGGAGAGTGTGTCACTCGGGAATGGTGAATAAGACCAAGCCACTGAATG 621
611 ATGGCATGGAGTCTGTGCACACTAGGCATGATAGAGAAACAAAGCTGAAGAACTC 670
622 AAAGTGTCTGGTGGCGGTACACACATTAACATTTGATTAATTTACAAGGAACATATCCA 681
671 AAGAAGGCTGGACTTACAGCATATAATCAAACTAGATACATCAAGAGAGATATATCCC 730
682 AAGGTGATCTCCAAGAAAGCTTTGATGATAGATGAAGAAATTTCAAAACCTTCAAGGA 741
731 AACATTAATACCAAGATCATATGATGATAGACTGCAGACTCTTGAGATGTCGGTGA 790
742 TCTGGATTAAGGCGATGCACAGGTGGTATTTCTTGGTCTTGGTGAGACCCAAAGGACCGT 801
791 GCTGGATTAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGGACCG 850
802 GTATCTTTCTTACACTTGGCCACAAATGATGATGATGATGATGATGATGATGATGATGAT 861
851 GTAGGGTGTGTGATACCTTACCTAGTACCTTGCCTTACACACCCAGAGAGCGTTCTTATTA 910
862 AGACTGTCTCCAACTCAAGGCGACGCCAAATGTATGAAGAGTGAAGAACCAAGTGA 921
911 GCATTTGCTGTAAAGGCAACCTCTTGGG-----ACCAAGAGCTGTAGAG 961
922 GTTATGAGATTTGTCAAGACCATTTGCTACCAAGATTTGGTCAATGCGCAAGCTTATATC 981
962 ATCTGGAAATGATCCGCATGATCGCCACTGCTCGGATCAGATGCCAAGGCAATGGTG 1021
982 AGATTGCTCAGGAAGATATACATGAAGAGGCGAAACAGGTGATGCTTCTCATGGCT 1041
1022 AGGCTTTTACAGGCGCGAGTACCGTCTTCGATGCCAAGCAAGCGCTGTCTTCTCGCT 1081
1042 GGTGTGAATCCATCTTCAAGGTGAAGAAATGCTCACAACATGTGTAAACGGCTGGAT 1101
1082 GGGGCCAACTCCATCTTTTGGCGGAGAACTTCTCAACCCGCAACCAACAGCTTTGAT 1141
1102 GAGGATAAGCCATTTGGCTAAATGGGGTCTGAACCAATGG 1144
1142 GCGGACCGGCGATTTCAAGATCCTTTGGCTGTATCCCAAGG 1184

RESULT 11
US-09-740-288A-17
; Sequence 17, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua

APPLICANT: Orozco, Emil
TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
FILE REFERENCE: BB1429 US NA
CURRENT APPLICATION NUMBER: US/09/740,288A
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/172929
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 1396
TYPE: DNA
ORGANISM: Hordeum vulgare
US-09-740-288A-17
Query Match 26.1%; Score 310.2; DB 9; Length 1396;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;
QY 202 AACACCTGGACCAAGAGAAATTAAGCTATATATGACACACCTCATGGACTTGATG 261
DB 213 AACGCTGGAGCGCGCGGAGATCCAGGCGCATCTACGACTCCCGCTCTCGACCTCTC 272
QY 262 CACTATGCTCAGGTGCAACACAGAGGTTTCCAAACCTTCAGAGGTTCAATTTGTCAC 321
DB 273 TTCCACGGGGCTCAAGTCCATAGGAATGTCCATAATTTAGAGAAGTGCACCAATGCACA 332
QY 322 CTTATGAATATCAAACTGTGTGTACCGGAGGACTGTAAGTACTGTGCCCAATCACAG 381
DB 333 CTTCTTTCAATTAAGACTGTGTGGTGCAGCGAAGATGTTCATCTGCCCACTGTTCA 392
QY 382 CGTTACAACACTGGTGTCAAGGCTGAAGAAATCATCAAGTTGATGAGGTTGAAGCT 441
DB 393 AGATACAGTACCGGATTGAAGCTGAAAAATTAATGAAGAAAGATGCCGCTCTAGAAGCA 452
QY 442 GCAAGGAGGCAAGGCCAAATGATCTACAAGTTCTGTATGGTGTCTGCTTGGAGAGAG 501
DB 453 GCTAAAAAGGCAAGGAGGCTGGAGACCCGATTTTGCATGGGAGCCGATGGAGAGAG 512
QY 502 ATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTCCATGAC 561
DB 513 ACATTTGGCAGGAAGAAACAACTTCAACAGATTTCTGATATGTCAAGGACATAAGAGGT 572
QY 562 CTTGGAATGGAGAGTTGTGTCACTCCCTGGGAATGGTTGATTAAGACCAAGCCACTGAATG 621
DB 573 ATGGGCATGGAGTCTGTTCACCTCCCTGGGCATGCTAGAGAAACAGCAAGCTGAAGAACTC 632
QY 622 AAAAGTGTGGTGTGACGGCGTACACCAATATGATGATGATGATGATGATGATGATGATGAT 681
DB 633 AAGAGGCTGGACTTACAGCTTATATCACTTATATCACTTATATCAAGAGAAATATTACCG 692
QY 682 AAGGTGATCTCCACAAAGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 741
DB 693 AACATTTCTTCAAGATCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 752
QY 742 TCTGGATTAAGGCAATGCAAGGTGTGTCTTGGTCTTGGTGAGACCCAAAGAGGACCGT 801
DB 753 GCTGGAATGAAGCTGTGTCTCAGGTGGAATTAATTTGGTCTTGGAGAGCGGAGGAGACCGT 812
QY 802 GTATCTTCTCTACACCTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGAT 861
DB 813 GTAGGCTGTGTGCATACACTGCGCCACTTTTGCACCAACACCCCAAGAGAGTGTTCCTATCAAT 872
QY 862 AGACTGTCTCCCAATCAAGGCGCACGCAATGTATGAAGAGTGAAGAAACCAAGCAAGTTGAA 921
DB 873 GCATTTGCTGTCAAGGCGACCGCTCTTCAAG-----ATCAGAGGCTGTAGAG 923
QY 922 GTTATGAGATTTGTCAAGAACCAATTTGCTACTGCAAGATTTGGTCAATGCCAACGCTTATATC 981
DB 924 ATATGGGAATATGATCGCATGATTTGCCAGGCTCGGATTTGTGATGATGATGATGATGATGAT 983
QY 982 AGATTGGCTGAGGAAGATATATCAATGAAGAGGCAAGAGTGTGATGATGATGATGATGATGATGAT 1041

Db 984 AGACTTTCCGAGGCGAGTACGGTTCTCCATGCCAGCAAGCTCTCTGCTTCTGCT 1043
Qy 1042 GGTGTAAATCCATCTTTACAGGTAAGAAAATGCTCAACAATGTGTAAAGCGCTGGGAT 1101
Db 1044 GGGCCAACTCGATCTTTCCGCGTGAAAAGCTCTCGACAACTGCAAAACAGGACTTTTGTAT 1103
Qy 1102 GAGGATAAGCCATGTTGGCTTAATGGGCTGTAACCAATGG 1144
Db 1104 GCGGACCAAGCAATGTTCAAGATCCTTGGCTGATTCCTCCCAAGG 1146

RESULT 12
US-10-431-544-17
; Sequence 17, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431.544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-10-431-544-17

Query Match 26.1%; Score 310.2; DB 16; Length 1396;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 213 AACGACTGGACCGCCGCGAGATCCAGGCCATCTAGACTCCCGCTCTCGACCTCTC 272
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTCCAACTTCAGAGGTTCAATTTGTCAGCT 321
Db 273 TTCCACGGGGCTCAAGTCCATAGGAATGTCCATAAATTTAGAGAAGTGCAACAATGC 332
Qy 322 CTTTATGAATCAAACTGCTGTGTGTACCGAGGACTGTAAAGTACTGTGCTCCCAATCAC 381
Db 333 CTTCTTTCAATAAGACTGTGGGTGCGAGAGATTTTCATCTGCTCCCAAGTCTTCA 392
Qy 382 CGTTACAACACTGCTGTCAAGGCTGAAAGAAATCATCCAGTTGATGAGGTGATTGAAGCT 441
Db 393 AGATACAGTACCGGATTTGAAGCTGAAATTAATGAAGAAAGATGCCGTCTTAGAAGCA 452
Qy 442 GCAAGAGGCAAAAGGCAATGATCTCAAGGTTCTGTATGGGTGCTGCTTGGAGAGAG 501
Db 453 GCTTAAAGGCAAAAGGAGGCTGGGAGCACCCGATTTTGTATGGAGCGCATGGAGAGAG 512
Qy 502 ATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGAC 561
Db 513 ACAATTTGGCAGGAAACAAACTTTCAACAGATTTCTTGAATATGTCAAGGACATGAAGGT 572
Qy 562 CTTTGAATGGAGAGTTGTGTCCACCTGGGAATGGTTGATAAAGCAACCAAGCACTGAATTG 621
Db 573 ATGGGATGGAGTCTGTTGCCACCTGGGATGCTAGAGAAACAGCAAGCTGAAGAACTC 632
Qy 622 AAAAGTGTGGGTGACGGGTACAACTTCAACATTGATATTTCAAGGAACATCTTCA 681
Db 633 AAGAAGGCTGGACTTACAGCTTATAATCAATCAACCTAGATACATCAAGAGAAATATTACCG 692
Qy 682 AAGGTATCTCCACAGAGAGCTTTGATGATAGATTGAAACATTTCAAAACAGTTCAAGA 741
Db 693 AACATTTATTTCTACAAGATCGTATGATGATGATTAAGACTCTTCAGCATGTCCGTGAA 752

Qy 742 TCTGGATTAAAGGCATGCAAGGTGTAATTTCTTGGTCTTGGTGAAGAGGACCGT 801
Db 753 GCTGGAATAAGCGTCTGCTCAGGTGGAATTAATTTGGTCTTGGAGAGGCGGAGGACCGT 812
Qy 802 GTATCTTTCTCTACACCTTGGCCCAATGATGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
Db 813 GTAGGCTGTGTGATACACTGGCCATTTTGGCCAAACACACCCAGAGAGTGTTCCTATCAAT 872
Qy 862 AGACTGCTCCCAATCAAGGCGACGCCAAATGTATGAAGAAAGTTAAGAAACAGCAAGTTGAA 921
Db 873 GCATTGATTCTCTCAAGGCGACGCTCTTCAGG-----ATCAGAGCCTGTAGAG 923
Qy 922 GTTGATGAGATTGTCAGAACCATTTGCTACTGTGAAGATTGTCATGCAACAGTCTATTATC 981
Db 924 ATATGGGAAATGATCCGCATGATTGCCAGCGCTCGGATTTGTGATGCCAAAGGCAATGGT 983
Qy 982 AGATTGCTGCGAGGAAGATATCAATGAAGAGGCGAGACAGGTGATGTCTTCTCATGGCT 1041
Db 984 AGACTTTCCGCGAGGCGAGTACGGTTCTCCATGCCAGAGCAAGCTCTCTGCTTTCTTGT 1043
Qy 1042 GGTGTAATGCCATCTTTCACAGGTAAGAAAATGCTCACAACAATGTGTAAACGGCTGGAT 1101
Db 1044 GGGGCCAACTCGATCTTCCGCGGTGAAGAGCTCTGCAACTGCCAAACAGCACTTTGAT 1103
Qy 1102 GAGGATAAAGCCATGTTGGCTAAATGGGCTTGAAGCAATGG 1144
Db 1104 GCGGACCAAGCAATGTTCAAGATCCTTGGCTGATTCCCAAGG 1146

RESULT 13
US-09-740-288A-23
; Sequence 23, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740, 288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Zea mays
US-09-740-288A-23

Query Match 26.1%; Score 310.2; DB 9; Length 1439;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 235 AACGACTGGAGCGGCGCGAGATCCAGGCGTCTAGCATCTCCGCTCTCTCGACCTCTC 294
Qy 262 CACTATGCTCAGGTGCAACACAGAAAGGTTTCCAAACACCTTCAGAGGTTCAATTTGTGCACT 321
Db 295 TTTCACGGGCTCAGTCCACAGAAATGTCCATAAATTCAGAGAGTGCGCAATGCACA 354
Qy 322 CTTATGAATATCAAACTGCTGTTGTACCGAGGACTGTAAAGTACTGTGCGCAATCACAG 381
Db 355 CTTCTTTCAATCAAGACTGGTGGATGCAAGATTTGTTTACTGTCTCTCAGTCATCA 414
Qy 382 CGTTACAACACTGCTCAAGGCTGAAGAAATCAATCCCAAGTTGATGAGGTGATTGAAGCT 441
Db 415 AGATACAACACTGGATTGAAGGCCCAAAATTAATGAAGCAAAATATGCTGTCTTGGAGCA 474

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28968
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4767-018-H9_FLI
US-10-425-114-28968

Query Match 26.1%; Score 310.2; DB 18; Length 1486;
Best Local Similarity 59.0%; Pred. No. 4.9e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACCACTCATGGACTTGATG 261
Db |||||
Qy 241 AACGACTGGAGCGCGCCGAGATCCAGCGCTTACGACTCACCGCTCTCGACTCCTC 300
Db |||||
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTGCACT 321
Db |||||
Qy 301 TTTCACGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAAGTGCAGCAATGCACA 360
Db |||||
Qy 322 CTTATGAATATCAAACTGGTGTGTACCGAGACTGTAAAGTACTGTGCCCAATCACAG 381
Db |||||
Qy 361 CTTCTTTCAATCAAGACTGGTGGATGCACTGAAGATGTTCTTACTGTCTCAGTCATCA 420
Db |||||
Qy 382 CGTTACAACTGGTGTCAAGCTGAAAGAAATCAATCAAGTGTGATGAGGTGATTTGAAGCT 441
Db |||||
Qy 421 AGATACAACTGGATTTGAAGGCCCAAAATTTGATGAACAAATATGCTGTCTTGAAGCA 480
Db |||||
Qy 442 GCAAGAGGCAAGGCCAATGGATCTACAAGTTCTGTATGGTGTCTGTGGTGGAGAGAG 501
Db |||||
Qy 481 GCAAAAAGGCAAAAGAGTCTGGAGCACCCGTTTTCATGGGAGCTGCGATGGAGAGAA 540
Db |||||
Qy 502 ATGAAGGTGAAAGTCAAACTTCAAGAAATCAAGAGATGATCACTGCTGTCTCCATGAC 561
Db |||||
Qy 541 ACCATTGGCAGGAATCAAACTTCAACAGATCTTGAATATGTCAAGGAAATTAAGGGGT 600
Db |||||
Qy 562 CTTGGAATGGAGAGTTGTGTCACTGGGAATGGTGTGATTAAGACCAAGCCACTGGAATTG 621
Db |||||
Qy 601 ATGGCATGGAGGTCTGTGACACTAGGCATGATAGAGAAACAAAGCTGAAGNACTC 660
Db |||||
Qy 622 AAGAGTCTGGTGTGACGGGTACAACTAACTTATGATCTTCAAGGAACTATCCA 681
Db |||||
Qy 661 AAGAAGCTGGACTTACAGCATATAATCAATCAACCTAGATACATCAAGAGAGTATATCCC 720
Db |||||
Qy 682 AAGTGATCTCCAAAGAGCTTTGATGATAGATTGAATCAATCAAAACGTTCAAGGA 741
Db |||||
Qy 721 AACATTATTACCAAGATCATATGATGATAGACTGTGAGACTCTTGAAGCATGTCGGTGA 780
Db |||||
Qy 742 TCTGGATTAAGGCATGCACAGGTGGTATTTCTGGTCTTGGTGGAGACCAAGAGGACCGT 801
Db |||||
Qy 781 GCTGGAATAGCATCTGCTCAGGTGGAATCATTTGTTCTTGGTGAAGCAGAGGAGCCGG 840
Db |||||
Qy 802 GTATCTTTCTTACCTTTGGCCACAATGGAATCAGATCCAGATCCAGATCTTTCCAATCAAC 861
Db |||||
Qy 841 GTAGGGTTGTTGCATACCTTAGCTACCTTGCCTACACACCCAGAGAGCGTTCTTATTAAT 900
Db |||||
Qy 862 AGACTGTCCCAATCAAGGCCACGCCAATGTATGAAGAAGTTAAGACACAGCAAGTTGAA 921
Db |||||
Qy 901 GCATTGGTTGCTGTAAAGGCAACACCTCTTTGAGG-----ACCAAGAGCCTGTAGAG 951
Db |||||
Qy 922 GTTGATGAGATTGTGCAAGACCATTTGCTACTGCAAGATTGGTCATGCCAACGTTCTATTATC 981
Db |||||

Search completed: September 17, 2005, 21:28:17
Job time : 920 secs

Db 952 ATCTGGGAAATGATCCGCATGATGCCCACTCTCGGATCAGCATGCCAAAGCAATGGTG 1011
Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAAGAGGCAGAACAGGTGATGCTTCTCATGGCT 1041
Db 1012 AGGCTTTTCAGCAGGCCGAGTACGGTTCTCGATGCCAGAAACAGCGCTGTGCTTCTCGCT 1071
Qy 1042 GGTGTAATGCTCTTTCACAGGTAAGAAATGCTCACAACAATGTGTAACGGCTGGGAT 1101
Db 1072 GGGGCCAACTCCATCTTTGCGCGGAGAACTTCTCACAACCGCAACACGACTTTGAT 1131
Qy 1102 GAGGATAAGCCATGTTGGCTTAAATGGGGTCTGAAACCAATGG 1144
Db 1132 GCGGACCAGGCGATGTTCAAGATCCTTTGGCCTGATCCCAAGG 1174

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 15:54:19 ; Search time 4431 Seconds
(without alignments)
10205.451 Million cell updates/sec

Title: US-09-752-957F-1
Perfect score: 1188
Sequence: 1 atgtcgcttatattgactgc.....aggtgcattcggtgttga 1188

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	333.4	28.1	969	CNS06HEG	AL398894 T7 end of
2	330.4	27.8	985	CNS06234	AL421814 T3 end of
3	297	25.0	1301	CNS0A9XF	EX819210 Arabidops
4	296.2	24.9	955	CNS0735Q	AL427092 clone BA0
5	294.4	24.8	1106	CNS0A9QQ	EX818983 Arabidops
6	292	24.6	1412	CY109523	AY109523 Zea mays
7	291	24.5	969	CO004857	CO004857 EST793192
8	290.8	24.5	876	CO084519	CO084519 GR_Ea01D
9	270	22.7	872	CF822545	CF822545 EST69927
10	266	22.4	834	CF822544	CF822544 EST69926
11	262.4	22.1	829	CF7219376	CF7219376 CCABF1670
12	252.4	21.2	862	CF719376	CF719376 CCABF1670
13	251.6	21.2	781	CD852378	CD852378 DH0ALL272
14	251.4	21.2	770	BU063672	BU063672 Fgr_3 G19
15	250.2	21.1	671	CAY81203	CAY81203 017F07AF
16	248.6	20.9	780	CF680028	CF680028 CCAD710TF
17	242	20.4	729	BU059880	BU059880 Fgr-C 0 G
18	240.6	20.3	811	CF448691	CF448691 EST685036
19	239.4	20.2	865	CF682137	CF682137 CCAGC25TR
20	239.4	20.2	948	CF721975	CF721975 CCABF1670
21	239.4	20.1	635	BU634775	BU634775 008D07 In
22	237.2	20.0	822	CO147058	CO147058 EST822111
23	237	19.9	825	CF690143	CF690143 CCAB241TF
24	234.4	19.7	721	BU064578	BU064578 Fgr_5_D19

25	233.4	19.6	509	8	AZ931212	AZ931212 474.dhz63
26	233	19.6	644	1	AV9233483	AV9233483 AV9233483
27	232.4	19.6	694	6	CA069980	CA069980 SCSGAD100
28	231.4	19.5	703	5	BU062717	BU062717 Fgr_2_C13
29	229.2	19.3	646	5	BU062716	BU062716 Fgr_2_C13
30	228.6	19.2	975	9	CNS06E0E	AL394500 T3 end of
31	227.4	19.1	596	1	AV9233288	AV9233288 AV9233288
32	226.6	19.1	730	1	AI960629	AI960629 sc87b11.Y
33	225.2	19.0	707	7	CF682132	CF682132 CCAGC25TF
34	225.2	19.0	826	7	CF719378	CF719378 CCAEF16TR
35	219.8	18.5	891	7	CO117806	CO117806 GR_EB01N
36	217.2	18.3	585	1	AI775752	AI775752 EST256852
37	216.2	18.2	730	5	BU025073	BU025073 QHF7113.Y
38	206.6	17.4	766	7	CN145395	CN145395 WOUND1_28
39	206	17.3	724	6	CA583129	CA583129 EST002804
40	204.2	17.2	796	4	BG582087	BG582087 EST483825
41	204	17.2	831	6	CD380074	CD380074 PTM05243
42	203.2	17.1	604	7	CO136883	CO136883 EST831554
43	201.2	16.9	588	2	AW980387	AW980387 EST391540
44	199.4	16.8	681	7	CF449199	CF449199 EST68544
45	199.2	16.8	814	6	CB293284	CB293284 UCRCS01_0

ALIGNMENTS

CNS06HEG 969 bp DNA linear GSS 30-NOV-2001
T7 end of clone AS0AA011H02 of library AS0AA from strain CLIB 533
of Saccharomyces bayanus, genomic survey sequence.
AL398894
AL398894.1 GI:12152774
GSS.

ORGANISM
Saccharomyces bayanus

REFERENCE
1 (bases 1 to 969)

AUTHORS

Solet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
Boulet, F., Bon, E., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durand, P., Lepoint, A., Llorente, B.,
Malpuech, A., Neveu, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.

TITLE

Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies

JOURNAL

FEMS Lett. 487 (1), 3-12 (2000)

MEDLINE

20584711

REFERENCE

2 (bases 1 to 969)

AUTHORS

Bon, E., Neveu, C., Casaregola, S., Artiguenave, F., Wincker, P.,
Aigle, M. and Durand, P.

TITLE

Genomic exploration of the hemiascomycetous yeasts: 5.
Saccharomyces bayanus var. uvarum

JOURNAL

FEMS Lett. 487 (1), 37-41 (2000)

MEDLINE

20584715

REFERENCE

3 (bases 1 to 969)

AUTHORS

Genoscope.

Direct Submission

JOURNAL

Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
secre@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of


```

the other extremity of this insert.
FEATURES
    source
        1..969
            /organism="Saccharomyces bayanus"
            /mol_type="genomic DNA"
            /strain="CLIB 533"
            /variety="uvaxum"
            /db_xref="taxon:4931"
            /clone="AS0AA011H02"
            /clone_lib="AS0AA"
            /note="end : 17"
            <2..>658
            /note="similar to Saccharomyces cerevisiae ORF YGR286c [
            BIO2 ; biotin synthetase ]"
            /evidence=not_experimental

misc_feature
    <2..>658
    /note="similar to Saccharomyces cerevisiae ORF YGR286c [
    BIO2 ; biotin synthetase ]"
    /evidence=not_experimental

ORIGIN
Query Match      28.1%; Score 333.4; DB 9; Length 969;
Best Local Similarity 71.0%; Pred. NO. 4.4e-83;
Matches 458; Conservative 1; Mismatches 177; Indels 9; Gaps 1;

Qy 522 CTTGAAGAAATCAAGAGATGATCCTGCTCCATGACCTTGGATGGAGAGTTGTGT 581
Db 4 CATGAAGAGAAATGAGGAGATGATTTCAAAGTGAACGACATGGGCTGGAAACATGTGT 63

Qy 582 CACCTGGGAATGGTTGATAAGACCAAGCCACTGAATTTGAAAGTGTCTGGTTGACGGC 641
Db 64 TACTTTGGGTATGGTCGACCAAGACCAGCAAGCAGCTAAAGATGACAGTGTGACCGC 123

Qy 642 GTACAACCATTAACATTTGATCTTCAAGGAACACTATCAAGAGTGATCTCCACAAGAG 701
Db 124 ATACAACCATTAACATTTGATCTTCAAGGAACACTATCAAGAGTGATCTCCACCAGAAC 183

Qy 702 CTTTGATGATGATTAAGAAACATTCAAAAGCGTTCAAGGATCTGATTAAGGATCGAC 761
Db 184 CTATGACATGATGATGACGACCATTAAGAACGTTCCAAAGAGTCCGGRATTAAGAGCTGCAC 243

Qy 762 AGGTGGTATTCTTGCTTTGGTGACCCAAAGAGGACCGTGTATCTTTCTCTACACCTT 821
Db 244 TGGTGGTATCTTAGTCTAGGTGAAGCGAAGATGACCAATGCTGTTTCATTATACAGTT 303

Qy 822 GGCCACATGGATCAGCATCCAGAGTCTTCTCCAAATCAACAGACTGGTCCCAATCAAGG 881
Db 304 GTCCAAATGATGACCCACATCCAGAGTCCCTACCAATCAACAGACTGTTGTCTATCAAGG 363

Qy 882 CACGCCAATGATGAGAGAGTTAAGAAC-----AAGCAAGTTGAAGTTGATGATGAT 932
Db 364 GACTCCAAATGGCTGAGGAAGTTCGCGACCCGCAAGTAAGAAAGTTGCAATTCGATGAAT 423

Qy 933 TGTCAGAACCATTTGCTACTGCAAGATTTGGTTCATGCCAAGCTCTATTATCAGATTGGCTGC 992
Db 424 CTTAAGAACTATCGCCACAGCAGCAAGTTGTCTATGCTAAGGCGATCAAGACTGCTGC 483

Qy 993 AGGAAGATATCAATTAAGAGAGGCGAGAACAGGTGATGTCTTCATGCTGCTGTTGTAATGC 1052
Db 484 TGGTGGTTACAAATGAAGGAGACCGAACAAATTTGCTGCTTCATGCTGCTGTTGTAACAG 543

Qy 1053 CATCTTCAGGTAGAAAATGCTCACAACATGTGTAAACGGCTGGATGAGGATTAAGC 1112
Db 544 TATTTTCACCGGTAGAAAATGCTTACAACTATGTGTAAACGGTGGGATGAAGACAAAGC 603

Qy 1113 CATGTTGGCTAAATGGGGTCTGAAACCAATGGAGAGTTTCAAATA 1157
Db 604 AATGTTGGCTAAGTGGGGCTTGCAACCTATGGAAGCCTTCGATA 648

RESULT 2
CNS06234/c
LOCUS
DEFINITION
  T3 end of clone AY0AA015C04 of library AY0AA from strain CBS 6340
  of Kluyveromyces thermotolerans, genomic survey sequence.
ACCESSION
  AL421814
VERSION
  AL421814.1 GI:12205009

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
    Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
    Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
    de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
    Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
    Saurin,W., Tekala,P., Toffano-Nioche,C., Wesolowski-Louvel,M.,
    Wincker,P. and Weissenbach,J
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711
PUBMED
11152876
2 (bases 1 to 985)
Malpertuy,A., Llorente,B., Blandin,G., Artiguenave,F., Wincker,P.
and Dujon,B.
Genomic exploration of the hemiascomycetous yeasts: 10.
Kluyveromyces thermotolerans
FEBS Lett. 487 (1), 61-65 (2000)
20584720
PUBMED
11152885
3 (bases 1 to 985)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail :
seque@genoscope.cns.fr - Web : www.genoscope.cns.fr]
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. Hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
    source
        1..985
            /organism="Kluyveromyces thermotolerans"
            /mol_type="genomic DNA"
            /strain="CBS 6340"
            /db_xref="taxon:4916"
            /clone="AY0AA015C04"
            /clone_lib="AY0AA"
            /note="end : T3"
            complement(<1..>788)
            /note="similar to Saccharomyces cerevisiae ORF YGR286c [
            BIO2 ; biotin synthetase ]
            1 putative frameshift(s)"
            /evidence=not_experimental

misc_feature
    complement(<1..>788)
    /note="similar to Saccharomyces cerevisiae ORF YGR286c [
    BIO2 ; biotin synthetase ]
    1 putative frameshift(s)"
    /evidence=not_experimental

ORIGIN
Query Match      27.8%; Score 330.4; DB 9; Length 985;
Best Local Similarity 66.0%; Pred. No. 3.2e-82;
Matches 492; Conservative 1; Mismatches 252; Indels 1; Gaps 1;

Qy 167 CATTGGAATATGCAATTTGCTGAGTGAAGACACACAGTCAACCTGGACCAAGAAATTA 226
Db 796 CACTGCAGTTTCGSSSTGTGCTCGAGCAGCTTCGACAGTGTGCTCCAAAGAGCAGCTGC 737

Qy 227 AAGCTATATATGACACACCACTCATGGACTTGACTATGCTCAGGTGCAACACAGAA 286
Db 736 GAGAGATCTACACACACCGCTGCTGGAGCTTGTGCACATCGCGCAGCTGCAGACCGCC 677

Qy 287 GGTTCCAACACCTTCAGAGGTTCAATTTGTGCATCTTATCAATATCAAAACTGGTGT 346
Db 676 GCTGGCAGATCTCTACCAAGTGCAGCTGTGCACTCTTATGAACATCAAGAGCGGTGGT 617

```

```
QY 347 GTACCGAGGACTGTAAAGTACTGTGCCAATCAAGCGTTTCAACACACTGGTGTCAAGGCTG 406
Db 616 GTTCCCGAGGACTGTAAAGTACTGTCCGCGAGTCGTCGCGGTACTCGAGGGGGCTCAAGGCGG 557
QY 407 AAGAATATCAATCCAAAGTTGATGAGTGTGATCAAGCTGCAAAAGGAGGCAAAAGGCGCAATGGAT 466
Db 556 AGAAGCTGTATCCGTGATCGGTGCTCAAGAGAGCGCGAGAGGCGCAAGCAGACGAT 497
QY 467 CTACAAGGTTCTGTATGTTGGTGTCTTGGAGAGAGATGAAAGGTAGAAAATCAAACTTGA 526
Db 496 CCACCTGTTTCTGCTGGCGCGCTGCGGTGACATGAACGGCGCTAAGTCCGGS-TGC 438
QY 527 AGAATATCAAGAGATGATCACTGTCTGCATGACCTTGGATGGAGAGTGTGTCAACC 586
Db 437 GCAAGATCTCCGAGATGGTCAAGAAGTCAATGAATGGGCTTCGAGACGTGTGTACGC 378
QY 587 TGGGAATGTTGATAAGACCAAGCCACTGAATTTGAAAGTGTCTGCTTGCAGCGGTACA 646
Db 377 TGGGCATGTCAAGAGGATCAGCGAGCAGCTGAAGACGCGGAGCTCACCGGTACA 318
QY 647 ACCATAACATGATCTTACAAGGAACACTATCTCAAGGTGATCTCCACAAGAGCTTTG 706
Db 317 ACCATAACATCGACACTCGCGTGAGCACTTACCCCAAGGTGATCAGTCCAGAAGCTACG 258
QY 707 ATGATAGATGAAAACATTCAAAACGTTCAAGGATCTGGATTAAGAGGATGACAGGTG 766
Db 257 AGCAGCGTCTGAACAGATCAAGAACGTCAAAAGCGCGGCGCATCAAGGGGTGCACCGGTG 198
QY 767 GTATTCCTGCTGTGGTGAGACCCAGAGAGCGGTATCTTCTCTACACTTGGCCA 826
Db 197 GTATTCGGGGCTGGCGGAGACAGAGGAGGACCAATTTGGTTCGTCCACACTTGGCCA 138
QY 827 CAATGATCAGCATCAGAGTCTCTTCCCAATCAACAGACTGGTCCCAATCAAGGGCAGCG 886
Db 137 ACATGACACACACCGCGAGTCGTCGCAATCAACAGGCTGGTCCGCAATCAAGGGAACCC 78
QY 887 CAATGTATGAAGAAGTTAAGAACAG 912
Db 77 CCATGGAGGAGGATCTGCAGAAAAAG 52
```

RESULT 3

CNS0A9XF

LOCUS
DEFINITION

1301 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTFB57ZG05 of flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).

ACCESSION

BX819210

VERSION

GI:42469429

KEYWORDS

HTC; GSLT cDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1301)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1301)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
Life technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full

length

<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES

Source

Location/Qualifiers

1..1301

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTFB57ZG05"

/tissue_type="Flowers and buds"

/plasmid="pCMVSPORT 6"

complement(1..1301)

/gene="At2g43360"

ORIGIN

Query Match 25.0%; Score 297; DB 3; Length 1301;
Best Local Similarity 58.7%; Pred. No. 1.1e-72;
Matches 538; Conservative 0; Mismatches 370; Indels 9; Gaps 1;

QY 202 AACACTGGACCAAGAAAGAAATTAAGCTATATATGACACACCACTCATGACTTGATG 261
DB 245 AACGATGGAGTAGAGATGAATCAAGCTGTTTATGATTCCTCTTCTGACCTCCTC 304
QY 262 CACTATGCTAGGTGCAACACAGAGAGTTTCAACACCTTCCAGAGTTCAATTTGTGCAT 321
DB 305 TTCCATGGAGTTCAGGTTTCATAGACATGTTTCAATCTTCAGGAGGTACAAATGTACC 364
QY 322 CTTATGATATCAAACTGGTGGTTGTACCGAGGACTGTAACTACTGTGCCCCAATCACAG 381
DB 365 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGACTGTTCAATTTGCTCCTCAGTCTCG 424
QY 382 CGTTACACACTGGTGTCAAGGCTCAAGAAATCATCCAAGTTGTAGAGGTGATTTGAAGCT 441
DB 425 AGATATAGCATGGAGTTAAGGCACAAAGACTCATGTCTAAGGAGCTGTCAATTTGATGT 484
QY 442 GCAAGAGGCAAGGCCAATGGATCTACAAGTTCTGTATGGTGTCTGCTTTGGAGAGAG 501
DB 485 GCTAAGAGGCAAGAAAGAGCTGGGAGCACACGTTTTTGCATGGTGTCTGTTGGCGAGAT 544
QY 502 ATGAAGGTAGAAAGTCAAACTTGAGAAATCAAGAGATGATCACTGCTGTCCATGAC 561
DB 545 ACAATTGGACGGAAACCAACTTTCAGCCAGATTTCTTGAATACATCAAAAGAAATAAGAGGC 604
QY 562 CTGGAAATGGAGAGTTGTGTCACTCGGAATGGTTGATTAAGACCAAGCCACTGAATTTG 621
DB 605 ATGGGATGGAGTTTGTGTGCACCTTAGGCATGATTGAGAAACAAAGCACTAGAGCTA 664
QY 622 AAAAGTGTGGTTGACCGGTACAAACCATTAACATTTGATCTTACAAGGAACACTATCCA 681
DB 665 AAGAAGGCTGCGCTCACTGCTTATAACCAATCTTTGATCTTCAAGAGAGTACTACCCA 724
QY 682 AAGTGATCTCCACAGAGAGCTTTGATGATAGATTGAAAACATTCAAAACGTTCCAGGA 741
DB 725 AACGTCATCACTACTAGAAAGTTATGACGATCGCTTGAAGACTCTTTAGCCATGTTTCTGTAT 784
QY 742 TCTGATTTAAAGGATGCACAGGTGGTATTCTTGTCTTTGGTGAGACCAAGAGGACCGT 801
DB 785 GCTGAATCAACGTTTGTTCAGGAGGAATCATAGGCTTTGGTGAGGAGAGAGACAGA 844
QY 802 GTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
DB 845 ATAGTTTATACACAGCTGCGCAATCTTCTCTCACCCTGAGAGTGTTCCTCATTAAT 904
QY 862 AGACTGTGCCAATCAAGGGCAGCGCAATGTATGAAGAAAGTTAAGAACAAAGAGTTGAA 921
DB 905 GCTCTACTTGAGTGAAAGGTACTCTCTTGAAG-----ACCAAGAGCCAGTTGAG 955

AUTHORS Castelli V., Aury J.M., Jaillon O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.

TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1306)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BP/Full_length <http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES Location/Qualifiers

1..1306

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTFB342B08"

/tissue type="Flowers and buds"

/plasmid="pCMVSPORT 6"

complement(1..1306)

/gene="At2g94360"

ORIGIN

Query Match 24.8%; Score 294.4; DB 3; Length 1306;
Best Local Similarity 58.0%; Pred. No. 6e-72;
Matches 545; Conservative 0; Mismatches 386; Indels 9; Gaps 1;

QY 202 AACCTCGACCAAGAGAAATTAAGCTATATATGACACACCACTCGACTTGATG 261
DB 219 AAGGATTGGATGAGATGAATCAAGTCTGTATATGATTCCTCTTCTGACCTTC 278
QY 262 CACTATGCTCAGTGCACACAGAGGTTCCACACCTTCAGAGGTTCAATTGTCAC 321
DB 279 TTCCATGGAGCTCAGGTTTCATAGACATGTTTCATAACTTCAGGGAGGTACA 338
QY 322 CTTATCAATATCAAACTGTGTCTACGAGGACTGTAAGTACCTGCGCCCAATCA 381
DB 339 CTCCTCTCCATAAAGACTGTGGCTGTAGTGAAGCTGTTCAATTGCTCAGTCT 398
QY 382 CGTTACAACACTGGTGTCAAGGCTGAAGAATCATCAAGTTGATGAGGTGATGA 441
DB 399 AGATAAAGCACTGGAGTGAAGGCACAAAGACTCATGCTAAAGGACGCTGTT 458
QY 442 GCAAGGAGGCAAGCCCAATGGATCTCAAGGTTCTGTATGGTGTCTGCTGGAG 501
DB 459 GCTAAGAAGGCAAGAAAGCTGGAGCACACGTTTTTGCATGGGTGCTGCTGG 518
QY 502 ATGAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGCC 561
DB 519 ACAATTGACGGAACCAACTTCAGCCAGATCTTGAATACATCAAGAAATAG 578
QY 562 CTTGGAATGGAGATTGTTGTCACCTCGGAATGGTTGATAAAGACCAAGCCACT 621
DB 579 ATGGGATGGAATGTGCTGCACCTCAGGCATGATTGAGAAACACACAGCACT 638
QY 622 AAAAGTGTGGTTGACGGCTGACACCAATACATGATCTTCAAGGAACACTATCA 681

639 AAGAAGCGTGGCTCACTGCTTATAACCAACAATCTTGTATATCTTCAAGAGAGTACTACCCA 698

682 AAGGTGATCTCCACAAGAAGCTTTGTATGATAGATTGAAAACATTTCAAAAACGTTCAAGGA 741

699 AAGCTCATCACTACTAGAGATTATGACGATCGCTTGAACACTCTTAGCCATGTGCGTGAT 758

742 TCTGGATTAAAGCATGACAGGTGTATTCTTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801

759 GCTGGAATCAACGTTTGTGACGAGCAATCATAGGCGTTGGTGAGGAGGAGACAGAGA 818

802 GTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTTCAATCAAC 861

819 ATAGGTTTATTACACAGCTGGCAACACTTCTTCCACCTGAGAGTGTTCCTCAATTAAT 878

862 AGACTGGTCCCAATCAAGGCGCAGCCAATGTATGAAGAAGTTAAGAACAGCAAGTTGAA 921

879 GCTCTACTTGCAGTGAAGGCACTCTCTTGAAG-----ACCAGAAGCCAGTTGAG 929

922 GTTGATGATGTCTGAGAACCATTTGCTACTGCAAGATTGGTTCATGCCAAGCTCTATTATC 981

930 ATATGGGAGATGATCAGGATGATTGGAACCGCACGTTATTGTATGCGCAAAAGCGTGTG 989

982 AGATTGGCTGCGAGGAGATACAAATGAAGAGGCGACAGGTGATGCTTCATGGCT 1041

990 ACACCTCTGCTGTAGAGTCCGGTTCTCAATGTCGGAACAAGCTCTCTGTTCTTGGCT 1049

1042 GGTGTAAATGCCATCTTCACAGGTAAAGAAATGCTCACAACAATGTGTAACGCTGGAT 1101

1050 GGTGCATATCTTATCTTCACCGGAGAGAGCTTTTAAACACACACATGATGTTTGAC 1109

1102 GAGGATAAAGCCATGTTGGCTAAATGGGCTCTGAAACCAA 1141

1110 GCTGACACGCTCATGTTCAAGACATTAGGCGCTCATTCCTA 1149

RESULT 6

AY109523

LOCUS Zea mays CL765_1 mRNA 1412 bp linear HTC 17-OCT-2002

DEFINITION Zea mays CL765_1 mRNA sequence.

ACCESSION AY109523

VERSION AY109523.1 GI:21213276

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1412)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
Unpublished (2002)
Coe, E.H.

REFERENCE 2 (bases 1 to 1412)
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

Location/Qualifiers

1..1412

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:632870"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

FEATURES source

/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```
ORIGIN
Query Match      24.6%; Score 292; DB 3; Length 1412;
Best Local Similarity 55.2%; Pred. No. 2.9e-71;
Matches 521; Conservative 0; Mismatches 413; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTGTATG 261
    |||||
Db 242 AACGACTGGAGCGGCCGAGATCCAGCGCTACGACTACCGCTCTCTGACCTCTC 301
    |||||
Qy 262 CACTATGCTCAGGTGCAACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTGCACT 321
    |||||
Db 302 TTTCAGGGGCTCAGGTCACAGAAATGTCCATAAATTCAGAGAAAGTCAGCAATGCACA 361
    |||||
Qy 322 CTTATGAATATCAAACTGGTGTGTACCGAGGACTGTAACTACTGTGCCCAATCACAG 381
    |||||
Db 362 CTTCTTTCAATCAAGACTGGTGATGCTAGTGAAGATTGTTTACTGTCTCCTCAGTCATCA 421
    |||||
Qy 382 CGTTACAACTGTGTCAAGGCTGAAAGAAATCATCCAAGTTGTAGTGAAGTGAAGCT 441
    |||||
Db 422 AGATACACACTGGATTGAGGCCCNNTTGTGACAAAGATGCTGTCTNNNNNNNN 481
    |||||
Qy 442 GCAAGGAGGCAAGGCCAATGGATCTCAAGGTTCTGTATGGGTGCTGCTTGGAGAGAG 501
    |||||
Db 482 NNNNNNNNNNNNNNNNNCTGGGAGCACCCGNNNNNNNNNNNNNNNNNNNNNNNN 541
    |||||
Qy 502 ATGAAGGTAGAAAGCTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCTCCATGAC 561
    |||||
Db 542 ACCATTGGCAGGAATCAAACTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 601
    |||||
Qy 562 CTTGAAATGGAGAGTTGTGTCAACCTCGGAAATGGTTGATAAGACCAAGCCACTGAATTG 621
    |||||
Db 602 ATGGCATGGAGGTTCTGTGACACTAGGCATGATAGAGAAACACNAGGTGAAGACTC 661
    |||||
Qy 622 AAAAGTGTGGTTGACGCGGTACAAACCATTAACATTGTACTTCAAGGAAACACTATCCA 681
    |||||
Db 662 AAGAAGGCTGGACTTACAGCATATAATCATAACTAGATACATCAAGAGAGTATTATCCC 721
    |||||
Qy 682 AGGTGATCTCCACAGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 741
    |||||
Db 722 AACATTATTACCAAGATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 781
    |||||
Qy 742 TCTGATTAAAGGCATGACAGGTGGTATTCTTTGGTCTTTGGTGAGACCCCAAGAGGACCGT 801
    |||||
Db 782 GCTGGNATAGGCATCTGCTCAGGTGGAATCATTTGGTCTTTGGTGAAGCAGAGGAGCGG 841
    |||||
Qy 802 GTATCTTTCTCTACACCTTGGCCCAATAGGATTCAGATCCAGAGTCTTCTTCAATCAAC 861
    |||||
Db 842 GTAGGTTGTTGCATACCTAGCTACCTTGGCTTACACACCCAGAGAGCGTTCCTATTAT 901
    |||||
Qy 862 AGACTGTTCCCAATCAAGGCGACGCAATGTATGAAGAGTTAAGAACAGCAAGTTGAA 921
    |||||
Db 902 GCATTGGTTGCTGTAAGAGGCACACTCTTTAGG-----ACCAAGAGCTGTAGAG 952
    |||||
Qy 922 GTTGATGAGATTGTGAGAACCAATTTGCTACTGCAAGATTGGTTCATGCCAACGTTATTATC 981
    |||||
Db 953 ATCTGGGAATGATCCGATGATCGCACTGTCTCGATCACCATGCCAAGGCAATGTTG 1012
    |||||
Qy 982 AGATTGGTTCAGGAGATATACATGAAAGAGGAGGAGAACAGGATGATGTCTTCAATGGCT 1041
    |||||
Db 1013 AGGCTTTTCAGCAGGCGAGTACGGTTCTCGATGCCAGAACAGCGCTGTGCTTCTCGCT 1072
    |||||
Qy 1042 GGTGTAATGCCATCTTACAGGTGAAGAAATGCTCACAACAAATGTGTAACGGCTGGAT 1101
    |||||
Db 1073 GGGGCCAACTCCATCTTTGCGCGGAGAGAACTTCTCACAACCGCAAAACAGCACTTGAT 1132
    |||||
Qy 1102 GAGGATAAGCCATGTTGGCTAAATGGGGTCTGTGAAACCAATGG 1144
    |||||
```

```
Db 1133 GCGGACGAGCGATGTTCAAGATCCTTGGCTGATCCCAAGG 1175
    |||||
RESULT 7
LOCUS      CO004857
DEFINITION EST793192 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
            Kb Coccidioides posadasii cDNA clone CIEAL82 5' end, mRNA sequence.
ACCESSION  CO004857
VERSION     CO004857.1 GI:48511746
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
ORGANISM    Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1 (bases 1 to 969)
AUTHORS     Gardner, M.J. and Cole, G.T.
TITLE       Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL     Unpublished (2003)
COMMENT     Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: Gardner@tigr.org
Seq primer: M13 Reverse.
FEATURES    Location/Qualifiers
            1..969
            /organism="Coccidioides posadasii"
            /mol_type="mRNA"
            /strain="C735"
            /db_xref="taxon:199306"
            /clone="CIEAL82"
            /dev_stage="spherules"
            /lab_host="E. coli DH10B, T1 phage resistant"
            /clone_lib="Coccidioides posadasii spherule cDNA library,
            0.4 to 2.3 kb"
            /note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
            Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
            Kb"
ORIGIN
Query Match      24.5%; Score 291; DB 7; Length 969;
Best Local Similarity 60.6%; Pred. No. 5e-71;
Matches 499; Conservative 0; Mismatches 315; Indels 9; Gaps 1;

Qy 173 AATATGCAATTCAGTGAAGACACACAGTCAACACTGGACCAAGAAATTAAGCTA 232
    |||||
Db 156 AAGAGCGGCTCTCAGCCACAGCACCTCGGATGAACTGGACCCGAGAGAGATTGAGAAGA 215
    |||||
Qy 233 TATATGACACACCTCATGCACTTGTGCTGCTATGCTCAGGTGCAACACAGAGGTTCC 292
    |||||
Db 216 TCTAGATGCACTCTGACCGAATTGACATATGCTTCGGTACCGTTACCGTTCGCTTCC 275
    |||||
Qy 293 AACAACTTCAGAGGTTCAATTTGTGCACTCTTATGAATATCAAAACTGGTGTGTACCG 352
    |||||
Db 276 ATGACCCCGCGCCATCAATGTCATCTCTATGACATATAAAACCGGTGATGACGG 335
    |||||
Qy 353 AGGACTGTAACTGTGTCGCCCAATCAGACGCTTAAACACACTGGTGTCAAGGCTGAAGAA 412
    |||||
Db 336 AAGATTGCTGTACTGCGCCCAATCTTCCCGCTACAAGACCGGGCTGAAGCCACAAAC 395
    |||||
Qy 413 TCATCCAAAGTTGATGAGTGTGAGCTGCAAGAGGAGGCAAGCCCAATGATCTACAA 472
    |||||
Db 396 TAAGCTCCGTAGACTCCGCTGCTTGAGGCGCCGCGGTGCGCAAGAGAAATGGCAGCATC 455
    |||||
Qy 473 GGTTCGTATGGGTGCTGCTTGGAGAGAGATGAAGAGGTAGAAAGTCAAACTTTGAAGAAA 532
    |||||
Db 456 GATTCTGATGGGTGCGCCCTGCGGCACATGCGGCGCCGAAACCAATCTACGAACA 515
    |||||
Qy 533 TCAAAGAGATGATCACTGCTCTCCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGAA 592
    |||||
```

```

516 TCAAGAGATGGTTTCCGGAGTCGAGCTATGGGAATGGAGTTTGGCTTACTCTGGAA 575
593 TGGTTGATAAAGACCAAGCCACTGAAATTGAAAGTGTGGGTTCAGCGCGTACAAACCATTA 652
576 TGATTGAGCGGAGCAGCCCAAGGAGCTTAAAGAGGCTGGCTTCCAGCGCTTACATCACA 635
653 ACATTGATCTTACAAGGAACATATCAATCAAGGATCTCTCACAAGAGCTTTGATGATA 712
636 ATCTTGACAGCTCGAGAGAAATTCATCCATCCATCATCACACAGAGATCGTATGACGAGA 695
713 GATTGAAACATTCAAAAGCTTCAAGGATCTGGATTAAAGGATGACACAGGTGGTATTC 772
696 GACTGACAGCCCTGGAGCATGTCCGGATCCGGTATCAACGCTGTGTTCCGGTGGCATTC 755
773 TTGGTCTTGTGTAGACCAAGAGACCGTGTATCTTCTCTACACCTTGGCCACCAATGG 832
756 TGGGCTTAGGTGAACAGGACTCCGACCGCGTGGGTGATACACCGTCCGACCATTCG 815
833 ATCAGCATCCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGT 892
816 CATCTCAACCAGAGTCTTCCCGTGAACGCGTCTGTACCTATCAAGGGCAACCCCTG- 874
893 ATGAAGAGTTAAGAACCAAGATTGAAGTTGATGAGATGTCAGAACCAATTGCTACTG 952
875 -----GGCGAGCGGAAATGATCTCTTTTGACAAGGTTCTCCGAACCGTGGCTACGG 926
953 CAAGATTGGTCAATGCAACCTCTATTATCAGATTGGCTGCAGG 995
927 CTCGTATCTTGTGCCCATGTCCTATTCAGATTGGCCGCTGG 969

```

RESULT 8

CO084519

LOCUS

DEFINITION GR_Ea01D02.f GR_Ea Gossypium raimondii cDNA clone GR_Ea01D02 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 876)

Kim H., Yu Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

Plate: 01 row: D column: 02.

Location/Qualifiers

1..876

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ea01D02"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH10B"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

FEATURES

source

ORIGIN

Query Match 24.5%; Score 290.8; DB 7; Length 876;
 Best Local Similarity 59.5%; Pred. No. 5.5e-71;
 Matches 515; Conservative 0; Mismatches 342; Indels 9; Gaps 1;

QY 276 GCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTGCACTCTTATGAATATCAA 335
 DB 2 GGATCACAGATATGCTCATAAATTCAGAGAGGTGCTTAGAGTACTCTCTCTCAATCAA 61
 QY 336 AACTGCTGTTGTACCGAGGACTGTAAGTACTGTGCCCAATCACAGGTTTACAACACTGG 395
 DB 62 GACTGTGGATGTAGCCGAGTTGTTCTTATGGCCCTCAATCTCTAGGTACCATCTGG 121
 QY 396 TGTCAAGGCTGAAAGAAATCATCCAAAGTGTATGAGGTGATTGAAGCTGCAAGGAGGCAAA 455
 DB 122 CCTAAGCCCCCAACAGCTCATGACCAAGATGCCGTNTATGCAGCAGCTTAACAGGCTAA 181
 QY 456 GGCCTAATGGATCTCAAGGTTCTGTATGGGTCTGTCTGGAGAGAGATGAAGGTAGAAA 515
 DB 182 AGAGGCTGTGTAGTACACGCTTTTGTGATGGTGCAGCATGGAGAGACACAGTCCGAAGGAA 241
 QY 516 GTCAAACTTTGAAGAAATCAAGAGATGATCACTGCTGTCATGACCTTGGAAATGAGAG 575
 DB 242 AACTAACTTCAACCAAAATCTTGAATATGTAAAGAAATTAGGGATATGGGGATGGAGGT 301
 QY 576 TTGTGTCACTCCCTGGGAATGTTGATAAAGACCAAGCCACTGAAATTCGAAAAGTCTGGGTT 635
 DB 302 GTGTGCACTTTGGGATGCTGTGAGNAGCAGCAGCACTTGAACTTAAGAGGACAGTCT 361
 QY 636 GACGGCTCAACACCAATTAATGATACTTCAAGGAACTATCCAAAGGTGATCTCCAC 695
 DB 362 TACAGCTTCAACCAATATCTCGATCTTCAAAAGAAATATTATCCCAACATTATTACAAC 421
 QY 696 AAGAGCTTTGATGATAGATTGAAAACATTCAAAACCTTCAAGATCTGGATTAAGGC 755
 DB 422 AAGAACCTTATGATGAGCGGTTGGAACCCCTTCAACATGTCCGTGAAGCAGGAATTAATGT 481
 QY 756 ATGCACAGGTGGTATTTCTTGGTCTTGGTGAGACCCCAAGAGGACCGGTGATCTTTCTCTA 815
 DB 482 CTGTTCAAGGGGCAATATAGGCTTTGGAGAACAGTAGAGGACCGGTTGGTTTATTGCA 541
 QY 816 CACTTGGCCCACAATGGATCAGATCCAGAGTCTTTCCAATCAACAGATCGTCCCAAT 875
 DB 542 CACTTGGCTACACTCCCACTCCAGAGAGTGTCCCAATTAATGCTTGGTGGGT 601
 QY 876 CAAAGGACGCAATGATGAAGAAGTTAAGAACAGCAGTGAAGTTGATGAGATTGT 935
 DB 602 GAAAGGACACCTCTTCAAGATCAA-----AAGCCAGTGGAAATATGGAGATGAT 652
 QY 936 CAGAACCAATGCTACTGCAAGATTGCTCATGCCAAGCTTATATCAGATTGGCTGCAGG 995
 DB 653 TCGGATGATTGCTGCTGCTGCAATGATGCTGCAATCAATGGTCAAGTTGTGAGTGG 712
 QY 996 AAGATATACATGAAAGAGGACAGATGATGCTTCTTCAAGGTTGTTGTAATGCCAT 1055
 DB 713 CAGAGTTCGGTCTCTATGCCAGAGCAGCATATGTTTTCTTGGTGGCAATTTCTAT 772
 QY 1056 CTTCAAGGTAAAGAAATGCTCAACCAATGTGAACGCTGGGATGAGGATAAGGCCAT 1115
 DB 773 CTTCACTGGTGAAGAGCTATTGACAACTCTCTAACAATGATTATGATGCTGATCAAGTCAT 832
 QY 1116 GTTGGCTAAATGGGGTCTGAAACCAA 1141
 DB 833 GTTCAAAATCTTGGACTGATTTCCAA 858

RESULT 9

CF822545

LOCUS

DEFINITION

CF822545 872 bp mRNA linear EST 01-APR-2004
 EST699927 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb. Coccidioides posadasii cDNA clone CIDAG50 5' end, mRNA sequence.


```

ACCESSION      CF822545      GI:45928602
VERSION        EST
KEYWORDS      Coccidioides posadasii
SOURCE        Coccidioides posadasii
ORGANISM      Coccidioides posadasii
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE      1 (bases 1 to 872)
AUTHORS      Gardner,M.J. and Cole,G.T.
TITLE        Analysis of gene expression in Coccidioides posadasii mycelia and
               spherules via expressed sequence tags
JOURNAL        Unpublished (2003)
COMMENT      Other ESTs: EST699926
               Contact: Gardner MJ
               The Institute for Genomic Research
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301 838 3519
               Fax: 301 838 0208
               Email: gardner@tigr.org
               Seq primer: M13 Reverse.

FEATURES             source
   Location/Qualifiers
       1..872
           /organism="Coccidioides posadasii"
           /mol_type="mRNA"
           /strain="C735"
           /db_xref="taxon:199306"
           /clone="CIDAG50"
           /dev_stage="saprobic phase (mycelia)"
           /lab_host="E. coli DH10B, T1 phage resistant"
           /clone_lib="Coccidioides posadasii saprobic phase cDNA
               library, 2 to 4 kb"
           /note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
               Coccidioides posadasii saprobic phase cDNA library, size
               fractionated cDNA 2 to 4 kb"

ORIGIN
Query Match      22.7%; Score 270; DB 7; Length 872;
Best Local Similarity 61.3%; Pred. No. 4.5e-65;
Matches 435; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Qy 173 AATATGCAATTGTACGTGAAGACACAGTCACAGTCACTGACCTGGACCAAGAAATTAAGCTA 232
Db 163 AAGAGCGCGTCTCAGCCACAGCACCTCGGATGAAGTGGACCGGAGAGAGATTGAGAAGA 222
Qy 233 TATATGACACACCACTCATGTGACTTGATGATGCTCAGTGCAACACAGAGAGTTCC 292
Db 223 TCTAGATGCACCTCTGACCGAATTGACATATGCTTCGCTACGTTTACCGTCCCTTCC 282
Qy 293 AACCACTTCAGAGGTTCAATTGTGCACCTTTATGAATATCAAACTCGTGGTTGTACCG 352
Db 283 ATGACCCCGCGCCATTCAAATGTGCACCTCTCATGAACATAAAACCGGTGGATGCAGCG 342
Qy 353 AGGACTGTAACTGTGCCCAATCAGCGTTACACACTGGTGTCAAGGCTGAAGAA 412
Db 343 AAGATTGTCTGTACTGCGCCCAATCTCCCGTACAAGACCGGGCTGAAGGCCACAAAC 402
Qy 413 TCATCCAAAGTTGATGAGGTGATTGAAGCTGCAAGAGGCAAGGCCCAATGGATCTCAA 472
Db 403 TAAGCTCGTAGACTCCGTGCTTGAGGCGCCAGGGTGCACAGAGAAATGGCAGCATC 462
Qy 473 GGTTCGTATGGGTGCTGCTGGAGAGAGATGAAGAGGTAGAAGTCAAACTTTGAAGAAAA 532
Db 463 GATTCTGATGGTGGCGCTTGGCGGACATGCGAGGCGGAAACCAATCTACGAACA 522
Qy 533 TCAAAGAGATATCATCTGCTTCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGAA 592
Db 523 TCAAAGAGATGGTTTCGGAGTGGAGCTATGGGAATGGAAATTTGCGTTACTCTGGAA 582
Qy 593 TGGTTGATAAGAACCAAGCCACTGAATTGAAGTGTGGTGGTGTGACGCGGTACACACATA 652
Db 583 TGATTGCGGCGAGCGGCCAAGAGCTTAAAGACGCTGGCTTACCGCTTACATACACA 642
Qy 653 ACATTGATACTTACAAGGAACACTATCCAAAGGTTGATCTCCACAAGAGCTTTTGATGATA 712

Db 643 ATCTTGACAGTCGAGAGAATTCTTACCCATCATCACACAGAGATCGTATGACGAGA 702
Qy 713 GATTGAAACATTCAAAAAGTTCAAGGATCTGAAGGATTAAGGCATGACACAGTGTGTATTC 772
Db 703 GACTGCAGACCTGGAGCATGTCCGCGATGCGGATATCAAGCTGTGTGGGTGGCATTC 762
Qy 773 TTGGTCTTTGGTGAGACCCCAAGAGGACCGTGTATCTTTCTCTACACCTTGGCCCAATGG 832
Db 763 TGGGCTTAGGTGAACAGGACTCCGACCGGTGCGGTTGATACACACCGTCCGACATTGC 822
Qy 833 ATCAGATTCAGACTCTCTTCCATCAACAGACTGGTCCCAATCAAGGGC 882
Db 823 CATCTCACCCAGAGTCTTTCCCGTGAAACGCGCTCGTACCTATCAAGGGC 872

RESULT 10
CF822544/c
LOCUS
DEFINITION  CF822544      834 bp      mRNA      linear      EST 01-APR-2004
               EST699926 Coccidioides posadasii saprobic phase cDNA library, 2 to
               4 kb Coccidioides posadasii cDNA clone CIDAG50 3' end, mRNA
               sequence.
ACCESSION  CF822544      GI:45928601
VERSION    CF822544
KEYWORDS   EST.
SOURCE     Coccidioides posadasii
ORGANISM   Coccidioides posadasii
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE  1 (bases 1 to 834)
AUTHORS    Gardner,M.J. and Cole,G.T.
TITLE      Analysis of gene expression in Coccidioides posadasii mycelia and
               spherules via expressed sequence tags
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST699927
               Contact: Gardner MJ
               The Institute for Genomic Research
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301 838 3519
               Fax: 301 838 0208
               Email: gardner@tigr.org
               Location/Qualifiers
                   1..834
                       /organism="Coccidioides posadasii"
                       /mol_type="mRNA"
                       /strain="C735"
                       /db_xref="taxon:199306"
                       /dev_stage="saprobic phase (mycelia)"
                       /lab_host="E. coli DH10B, T1 phage resistant"
                       /clone_lib="Coccidioides posadasii saprobic phase cDNA
                           library, 2 to 4 kb"
                       /note="vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
                           Coccidioides posadasii saprobic phase cDNA library, size
                           fractionated cDNA 2 to 4 kb"

ORIGIN
Query Match      22.4%; Score 266; DB 7; Length 834;
Best Local Similarity 61.6%; Pred. No. 6.1e-64;
Matches 447; Conservative 0; Mismatches 270; Indels 9; Gaps 1;

Qy 430 GTGATTGAAGTGCACAAAGGAGGCAAGGCCAATGGATCTACAGGTTCTGTATGGGTGCT 489
Db 827 GTGCTTGAGGCGCGCCAGGTCGCCAAGAGAAATGGCAGCACTCGATTCTGATGGGTGCC 768
Qy 490 GCTTGGAGAGAGATGAAGAGGTAGAAAGTCAAACTTGAAGAAAATCAAGAGATGATCACT 549
Db 767 GCCTGGCGGACATGCGAGGCGGAAAAACCAATCTACAAAAACATCAAGAGATGGTTTCC 708
Qy 550 GCTGTCCATGACCTTGAATGGAGAGTTGTGTACCCCTGGGAATGGTTGATAAAGACCAA 609
Db 707 GGAGTGCAGCTATGGGAATGGAAATTTGCGTTACTCTGGGAATGATTCGCGGCGAGCAG 648

```

```
QY 610 GCCACTGAATTGAAAGTCTGGGTTGACGGCTGACCAACCATTAACATTGATCTTACAG 669
Db 647 GCCAGAGGCTTAAGACGCTGGCCCTACCGCTTACCAATCACAACTTTCACACGTCGAGA 588
QY 670 GAACACTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTGAAAACATTCAAA 729
Db 587 GAATTCCTACCCATCCATCATCACACGAGATCGTATGACGAGAGACTGCAGACCTGGAG 528
QY 730 AAGCTTCAAGGATCTGGATTAAAGGATGACAGGTGGTATTCTTGGTCTTTGGTGAGACC 789
Db 527 CATGTCGGCGATCCGGTATCAACGTCGTTCGGGTGGCAATCTGGGCTTAGGTGAACAG 468
QY 790 CAGAGGACCGGTATCTTCTCTACACTTGGCCACATGATGATGATCAGATCCAGATCT 849
Db 467 GACTCGGACCGGCTCGGGTTGATACACACCGTCCGACATTTGCCATCTCACCAGAGTCT 408
QY 850 CTTTCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGATCAAGAAGTTAAGAAC 909
Db 407 TTCCCGGTGAAACCGGCTCGTACCTATCAAGGGCAGCCCGCTGGCGAGCGGAAATGATC 348
QY 910 AAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTGGTCAATGCCA 969
Db 347 TCTT-----TTGACAAGGTTCTCCGAACCGTGGCTACGGCTCGTATCGTTCTGCCC 297
QY 970 AGCTCTATTATCAGATTGGCTGCAGGAAGATATACATGAAGAGGCAGAACAGGTGATG 1029
Db 296 ATGTCCATTTGTCAGATTGGCCCGCTGGCCGTTATCTCCATGACCGAAGAGCAACAGGTACC 237
QY 1030 TCGTTTCATGGCTGGTTGTAATGCACTTTTACAGGTAAGAAATGCTCACAACATGTCT 1089
Db 236 TCGTTTCAGGCTGGTGGCAATGGATCTTACAGGAGAGAGATGTTAATCCAGCTGC 177
QY 1090 AACGGCTGGATGAGGATAAAGCCATTTGGCTTAAATGGGCTGTAACCAATGGAGAGT 1149
Db 176 TCCGGTGGATGAGGATAAAGATATTTCGATAAGTGGGATATTACCCCATGCGACCT 117
QY 1150 TTCAAA 1155
Db 116 TTTGAA 111
```

```
RESULT 11
CF721966/c
LOCUS
DEFINITION
CF721966
VERSION
KEYWORDS
SOURCE
```

```
ORGANISM
Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
```

```
REFERENCE
1 (bases 1 to 829)
Loftus,B.
End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
Unpublished (2003)
Other_ESTs: CCA1517R
Contact: Brendan Loftus
TIGR
```

```
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: cryptotigr.org
Seq primer: TF.
```

```
FEATURES
source
1..829
Location/Qualifiers
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
```

```
ORIGIN
```

```
Query Match 22.1%; Score 262.4; DB 7; Length 829;
Best Local Similarity 60.6%; Pred. No. 6.4e-63;
Matches 453; Conservative 0; Mismatches 286; Indels 9; Gaps 1;

QY 413 TCATCCAAAGTTGATCAGGTGATTGAAGCTGCAAGAGGAGCAAGGCAATGATCTACAA 472
Db 821 TCGTCAGCATTTGAGCGCTGTGCTTGAAGCTGCTCGTCAGCCAGGCCACGCTCCACTA 762
QY 473 GGTTCGTGATGGGTGCTGCTTGGAGAGAGATGAAAAGGTAGAAAGTCAAACTTTGAAGAAA 532
Db 761 GATTCGTGATGGGTGCTGCTTGGAGAGACTTTGGCCGGTAGAAAGAGTGGATTTCGAAAAG 702
QY 533 TCAGAGAGATGATCACTGCTGTCCATGACCTTGGATGGAGAGTGTGTCAACCTGGGAA 592
Db 701 TCTTGGAAATGGTTAAAGAGGTTCAGGGGTATGGGAATGGAGGTCTGCACGACTTTGGGTA 642
QY 593 TGGTTGATAAAGACCACGACCTGAATTTGAAAAGTGCTGGGTGAGCGGCTACCAACATA 652
Db 641 TGCTCTCTCTGAGCAGCTATCCGGCTCAGGAAGCTGGTCTCAGCGCTATTAACACA 582
QY 653 ACATTGATCTTCAAGGAACAATATCAAGAGTGTATCCCAAGAGAGCTTTGATGATA 712
Db 581 ACCTTGACATCTCTCGAATTTCTATCCGAGGTGTCACTCAGATCATACGATGATC 522
QY 713 GATTGAAAACATTCAAAACGTTCAAGGATCTGGATTAAAGGCATGCAAGGTGGTATTC 772
Db 521 GATTGAGCACCATTTGCTGCCGTGGAGAGTGGTATCTCAGTCTGTCTTGGTGGCATCC 462
QY 773 TTGTTCTTGGTGAGACCCAGAGGACCGTGTATCTTTCTCTACACCTTGGGCAACAATGG 832
Db 461 TCGGTTTGGGAGAGGGGATGAGGATCGTTCGGATTGATCCAGAGGTTTCCAGGATGC 402
QY 833 ATCAGATCTCAGAGTCTTTCCAATCAACAGACTGGTCCCAATCAAGGGCAGCCAAATGT 892
Db 401 CCGAGCACCTTGAGTCTTTCCCGCTCAACACATTTGCTCCCTATCCTGCTACTCCCTCG 342
QY 893 ATGAGAGAGTTAAGAACAGCAAGTTGAGTTGATGAGATTGTTCAGAACCATTTGCTACTG 952
Db 341 AAGGC-----AAGAGCCTGTCAAGGTCCACACCGCTCTCCGAACCATCGCTACTG 291
QY 953 CAAGATTGGTTCATGCCAACGTCATTATTCAGATTGGCTGAGGAAGATATACAATGAAAG 1012
Db 290 CCGGTATCGTCTTCCAGAGCAATCATTCGTCTCGCGCTGGCGGACACGAATTCAGGG 231
QY 1013 AGGCAGAACAGGTGATGTGCTTCAATGAGTGGTGGTGTGTAATGCCATCTTCAAGGTAAAGAAA 1072
Db 230 AGACCGAACAGCCATGSCCTTTCATGSCGSGTGCCAAATGCCATCTTCACTGGCAGAGA 171
QY 1073 TGCTCACACAAATGTGTAAACGCTGGGATGAGGATGAAGCCATGTGTGGCTAAATGGGGTC 1132
Db 170 TGCTCACCAACCCCTTGTCTCCGGATGGGATGAGGCAAGGCTATGTGTGACAGGTGGGGTC 111
QY 1133 TGAACCAATGGAGAGTTTCAAAATACAA 1160
Db 110 TCAGGGCCAGAGAGAGCTTTGAGGACAA 83
```

```
RESULT 12
CF719376/c
LOCUS
DEFINITION
CF719376
ACCESSION
```

```
CF719376 862 bp mRNA linear EST 16-AUG-2004
CCAEF16TO C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCAEF16, mRNA sequence.
CF719376
```

```
/strain="JEC21"
/db_xref="taxon:40410"
/clone="CCA1517"
/clone_lib="C.neoformans strain JEC21"
/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"
```



```

VERSION      CF719376.1  GI:41573535
KEYWORDS     EST.
SOURCE       Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
              var. neoformans)
ORGANISM     Cryptococcus neoformans var. neoformans
              Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
              Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
              Filobasidiella.
REFERENCE    1 (bases 1 to 862)
AUTHORS      Loftus,B.
TITLE        End sequencing of clones from a full length enriched, normalized
              JEC21 cDNA library
JOURNAL      Unpublished (2003)
COMMENT      Other ESTs: CCAEP16TR
              Contact: Brendan Loftus
              TIGR
              Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
              Rockville, MD 20850, USA
              Tel: 301-838-3543
              Fax: 301-838-0208
              Email: cryptotigr.org
Seq primer: 1F.
              Location/Qualifiers
FEATURES     1..862
              /organism="Cryptococcus neoformans var. neoformans"
              /mol_type="mRNA"
              /strain="JEC21"
              /db_xref="taxon:40410"
              /clone="CCAEP16"
              /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full
              length, normalized library was prepared from a variety of
              conditions using RNA provided by Joseph Heitman and
              Jennifer Lodge"
ORIGIN
Query Match      21.2%; Score 252.4; DB 7; Length 862;
Best Local Similarity 60.5%; Pred. No. 4.5e-60;
Matches 437; Conservative 0; Mismatches 276; Indels 9; Gaps 1;

Qy 439  C T G C A A A G G G C C A A G C C A A T G G A T C T C A A G G T T C T G T A T G G G T G C T T G G A G A 498
      |||||
Db 862  C T G C T C T C A A G C C A A G C C A A C G G T C C A T A G A T T C T G A T T C G T G G T G C T T G G A G A 803
      |||||
Qy 499  G A G A T G A A G T G A G A A G T C A A A C T T G A A G A A A T C A A G A G A T G A C A C T G C T G T C C A T 558
      |||||
Db 802  G A C T T G C C G G T A G A A A G A G T G G A T T C G A A A A G A T C T T G A A A T G G T T A A A G A G G T C A G G 743
      |||||
Qy 559  G A C C T T G G A A T G G A G A G T T G T G T C A C C C T G G A A T G G T T G A T A A A G A C C A A G C C A C T G A A 618
      |||||
Db 742  G G T A T G G A A T G G A G G T C T G C A C G A C T T T G G T A T G C T C T C T G A C C A A G C T A T C C G G 683
      |||||
Qy 619  T T G A A A A G T T G G G T T G C A G C G G T A C A A C C A T A A C A T T G A T A C T T T A C A A G A A C A C T A T 678
      |||||
Db 682  C T C A A G A A G C T G G T C T C A G C G C C T A T A A C C A C A C C T T G A C A C T T C T C G A G A A T T C T A T 623
      |||||
Qy 679  C C A A G G T G A T C C A C A A G A G C T T T G A T G A T A G A T T G A A A C A T T C A A A A A G T T C A A 738
      |||||
Db 622  C C C G A G T T G C A C C T C A G A T C A T A C A T G A T G A T G A C A C A T T G T G C C G T G C G A 563
      |||||
Qy 739  G G A T C T G A A T T A A A G G C A T G C A C A G G T G G T A T T C T T G C T T G T G A G A C C A A G A G G A C 798
      |||||
Db 562  G A A G C T G T A T C A G T C T G T T G T G G C A T C C T C G G T T T G G A G A G C G G G A T G A G A T 503
      |||||
Qy 799  C G T G T A T C T T C C T T A C A C C T T G C C A A A T G G A T C A G A T C C A G A T C T C T T C C A A T C 858
      |||||
Db 502  C G T C G G A T T G A T C C A C A G G T T C C A G A T G C C C G A G C A C C C T G A G T C T T C C C G C T C 443
      |||||
Qy 859  A A C A G A C T G T C C C A A T C A A G G G C A G C C A A T G A T G A A G A A G T T A A G A C A A C A A G T T 918
      |||||
Db 442  A A C A C A T T G C C C T A T C C C T G G T A C T C C C C T C G A A G G C - - - - - A A C G A C G C T G T C 392
      |||||
Qy 919  G A A G T T G A T G A T T G T C A G A A C C A A T T G C T A C T G C A A G A T T G G T C A T G C C A A C G T C T A T T 978

```

```

Db 391  A A G G T C C A C A C C G T C C T C C G A A C C A T G C T A C T C C C G T A T C G T T C T T C C A A G A C C A T C 332
      |||||
Qy 979  A T C A G A T T G C T G C A G A A G A T A T A C A A T G A A A G G C A G A C A G G T G A T G T G C T T C A T G 1038
      |||||
Db 331  A T T C G T C T C G C C G T G C C G C A C A G A A T T C A G C A G A C C G A A C A A G C C A T G C C T T C A G 272
      |||||
Qy 1039  G C T G G T T G A T G A T G C A T C T T C A C A G G T A A G A A A A T G C T C A C A A C A A T G T T A A C G G C T G G 1098
      |||||
Db 271  G C C G G T C C A A T G C A T C T T C A C T G G C G A G A G A T G C T C A C C A C C C T T G C T C C G A T G G 212
      |||||
Qy 1099  G A T G A G A T A A A G C C A T G T T G G C T T A A A T G G G T C T G A A C C A A C C A A T G G A G A G T T C A A A T A C 1158
      |||||
Db 211  G A T G A G C A A A G C T A T G T T G G A C A G G T G G G T C T C A G G G C C A G A A G A A G C T T T G A G G A C 152
      |||||
Qy 1159  A A 1160
      ||
Db 151  A A 150

RESULT 13
CD852378
LOCUS      DHOALL272F09Z2M1 HaDevR5 Helianthus annuus cDNA clone HaDevR527F09,
DEFINITION mRNA sequence.
ACCESSION CD852378
VERSION    CD852378.1 GI:32536194
KEYWORDS   EST.
SOURCE     Helianthus annuus (common sunflower)
ORGANISM   Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE  1 (bases 1 to 781)
AUTHORS    Genoplate.
            Genoplate, a major partnership french program in plant genomes
TITLE      Unpublished (2003)
JOURNAL    Contact: Genoplate
COMMENT    Genoplate
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplate' (http://www.genoplate.com
            and http://genoplate-info.infobiogen.fr).
FEATURES   source
            1..781
            /organism="Helianthus annuus"
            /mol_type="mRNA"
            /culturvar="psc8"
            /db_xref="taxon:4232"
            /clone="HaDevR527F09"
            /tissue_type="4 days after self-pollination embryo"
            /clone_lib="HaDevR5"
ORIGIN
Query Match      21.2%; Score 251.6; DB 6; Length 781;
Best Local Similarity 58.5%; Pred. No. 7.3e-60;
Matches 462; Conservative 0; Mismatches 319; Indels 9; Gaps 1;

Qy 209  G G A C C A A G A A G A A A T T A A G C T A T A T A C A C A C C A C C A C T A T G A C T T G A T G C A C T A T G 268
      |||||
Db 1     G G A A C A A G A A G A G A T C A G T C G T T T A G A T T C T C T C T C T C T C T C T C A T C T C T G T T C A G 60
      |||||
Qy 269  C T C A G G T G C A A C A G A A G G T T C C A C A C A C C T T C A G A G G T T C A A T T G T G C A C T T T A T G A 328
      |||||
Db 61    G A G C T C A A G T T C A T A G A C A T G C T C A G A A T T T T A G G A G G T T C A G C A G T G A C G C T T C T G T 120
      |||||
Qy 329  A T A T C A A A C T G T G G T T G T A C C G A G A C T G T A A G T A C T G T G C C A T C A C A G G T T A C A 388
      |||||
Db 121  C T A T A A A G A C G G G T G G G T G C A G T G A A G A T T G T T A T T G C C C T C A A T C T T C T A G G T A T G 180
      |||||

```

```

QY 389 ACACCTGCTGCAAGGCTGAAGAATCATCCAGTTGATGAGGTGATTTGAAGCTGCAAGG 448
Db 181 ATACTGAGTGAAGCTCAAAAGCTCATGAACCAAGGATGCTGTTCTTTGAAGCTGCACAAA 240
QY 449 AGGCAAAAGGCAATGATCTACAAAGTTCTGATGAGGTGCTGCTGTTGAGAGAGATGAAG 508
Db 241 AGGCAAAAGGCTGAGCAGCTCGGTTTTCATGGGTGCTGCTGTTGAGAGAGACACATAG 300
QY 509 GTAGAAAGTCAAACTTTGAAGAAATCAAGAGATGATCACTGCTGTGCTCCATGACCTTGGAA 568
Db 301 GAAGGAAACCAACTTTAGTCAGATCCTTGAATATGTGAAGAGATGAAGGGGTATGGAA 360
QY 569 TGAAGAGTTGTGTCACTCCCTGGGAATGTTGATTAAGACCAAGCAGCTGAATTTGAAGATG 628
Db 361 TGAAGGTTTGTGTACGTTTGGGAATGATGAAGAAAGCAGCAGGCTTTAGAACTCAAGAAAG 420
QY 629 CTGCGTTGACGGGTGACCAACATAACATTGATCACTTTACAGGACACATATCCAAAGTGA 688
Db 421 CAGGCTTACAGCTTACATATCAATCAACCTTGTATACATCAAGAGAGTACTCCCTAATATCA 480
QY 689 TCTCCACAAGAGCTTTTGATGATAGATTGAAGAAATCAAAACAGTTTCAAGGATCTGGAT 748
Db 481 TTACAACAAGACATATGACNACGCTGGAACCATCAAGCATGTTCTGTGAAGCAGGA 540
QY 749 TAAAGGATGCAAGGTGATTTCTTGGTCTTGGTGAGACCCAGAGAGCCGCTGTATCTT 808
Db 541 TTAATGTTTGTCTGAGGAATAATAGGGCTTGGTGAAGCAGAAGAGGACAGGGTTGGGT 600
QY 809 TCTCTACACCTTGGCCACATGATGATGAGAAAGTGAAGAAAGTGAAGTGAAGTGAAG 868
Db 601 TGTGTATACATTAGCAACACTCCCTCACCACCGGAGAGTGTTCCTCATTAACGCACTTC 660
QY 869 TCCCAATCAAGGCGACGCAATGATGAAGAAAGTGAAGAAAGTGAAGTGAAGTGAAG 928
Db 661 TTGCAATTAAAGCACAACCTTTGAAGATCA-----AAGCAGATTGAGATGGG 711
QY 929 AGATTGTGAGAACCATTTGCTACTGCAAGATTGCTGATGCCAAGCTCTATTATCAGATTGG 988
Db 712 AGATGATACCGATGATTGCTACGCGCGCATCAGATGCCAAGCTATGGTGGATTTAT 771
QY 989 CTGCAAGGAG 998
Db 772 CTGACGGAG 781

```

```

RESULT 14
BU063672/c
LOCUS
DEFINITION
  Fgr_3_G19_T7 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
ACCESSION
  BU063672
VERSION
  BU063672.1
KEYWORDS
  EST.
SOURCE
  Gibberella zeae
  Gibberella zeae
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
  1 (bases 1 to 770)
AUTHORS
  Trail, F., Xu, J.-R., San Miguel, P., Hølgren, R.G. and Kistler, H.C.
TITLE
  Analysis of expressed sequence tags from Gibberella zeae (anamorph
  Fusarium graminearum)
JOURNAL
  Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE
  22508120
PUBMED
  12620255
COMMENT
  Contact: Frances Trail
  Department of Plant Biology
  Michigan State University
  East Lansing, MI 48824, USA
  Tel.: 517 432 2939
  Fax: 517 353 1926
  Email: trail@msu.edu
  Plate: 3 row: G column: 19.
  Location/Qualifiers

```

```

source
  1. 770
  /organism="Gibberella zeae"
  /mol_type="mRNA"
  /strain="NRRL 31084"
  /db_xref="taxon:5518"
  /clone_lib="Nitrogen-starved mycelia"
  /notes="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
ORIGIN
  Query Match      21.2%; Score 251.4; DB 5; Length 770;
  Best Local Similarity 62.7%; Pred. No. 8.3e-60;
  Matches 412; Conservative 0; Mismatches 236; Indels 9; Gaps 1;
QY 495 GAGAGAGATGAAGGTAGAAAGTCAAACTTTGAAGAAATCAAGAGATGATCACTGCTGT 554
Db 756 GAGGACATGCGCGGACGCAAGAACACCTCAAGACATCAAGCCCATGTCGAGGGCGT 697
QY 555 CMTAGACCTTGAATGGAGAGTTGTCTACCTCTGGGAATGTTGATTAAGACCAAGCCAC 614
Db 696 CAAGGCGATGGGAATGGAGGTCTCGCTAAACGCTGGGCATGATCGACGCGAGCAGCCAA 637
QY 615 TGAATTTGAAGTGTCTGGGTTCACGCGGTACACCATTAACATTTGATCACTTACAAGNACA 674
Db 636 GGAGCTCAAGGCGGAGGACTCACGCGGTACCAACCAATGTTGATCTAGCCGAGATT 577
QY 675 CTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTGAAACATTTCAAAAACGT 734
Db 576 TTACCCCAACGTCATCACGCGGAGCTACGACGAGCAGCTGCAGACACTGAGTCATGT 517
QY 735 TCAAGGATCTGAATTAAGGCAATGACAGGTGATTTCTTGGTCTTTGGTGAGACCCAAAGA 794
Db 516 CCGTGATGCTGGAATTAATGTCGCTCTGCTGGAATTTCTGGGTCTTTGGTGAGACGCGA 457
QY 795 GGACGCTGTATCTTCTCTACACTTGGCCACAATGGATCAGCATCAGAGTCTCTTCC 854
Db 456 GGACGCTGTGTTGTTGCTCCATACCGTCTCGACGCTCCCAAGCCACCGGAGAGCTTCCC 397
QY 855 AATCAACAGACTGTGCTCCCAATCAAGGCGACGCAATGATGAAGAAAGTTAAGAAACAGCA 914
Db 396 CGTGATGCTGCTGTTCTTATCAAGGTACACACTTGGTGACAGACCGATG----- 345
QY 915 AGTTGAAGTTGATGAGATTGTCAGAACCATTTCTACTCAAGATTGGTATGCTCAACAGCT 974
Db 344 -GTGGAGTTTACGAGCATGCTGCGACCATCGCGCGCGCGCATCATATGCGCGTCGAC 286
QY 975 TATTATCAGATTGGCTGAGGAAGATATACATGAAGAGGCGAGACAGGTATGCTT 1034
Db 285 CATCATCCGATTCGACGCGGTCGCAAGACCATGTTCAGAGGAGAAAGCAAGCGCTATGCTT 226
QY 1035 CATGCTGCTGTTGATGTCATCTTCAAGGTGAAGAAATGCTCAACAATGTTGTAACCG 1094
Db 225 CATGCGAGTGGCAACGCCAICTTTACGGGTGAAGAGATGCTGACACCGAGTGTANTGG 166
QY 1095 CTGGATGAGGATTAAGCCATGTTGGCTAAATGGGCTCTGAAACCAATGAGAGTTT 1151
Db 165 CTGGATGAGGATGCGGCCCATGTTTGGTCGATGGGCTCTGGAGGCCCATGAAGCTT 109
RESULT 15
CA781203
LOCUS
DEFINITION
  CA781203 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
ACCESSION
  CA781203
VERSION
  CA781203.1
KEYWORDS
  EST.
SOURCE
  Arabidopsis thaliana (thale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
  1 (bases 1 to 671)
REFERENCE
  1
AUTHORS
  Lundsgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.

```

